



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/45, 15/47, 15/40, C07K 14/12, 14/115, 14/135, C12N 7/04, A61K 39/165, 39/155, C07K 14/145, 14/08</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 98/13501</b> <b>(43) International Publication Date:</b> 2 April 1998 (02.04.98)
<b>(21) International Application Number:</b> PCT/US97/16718 <b>(22) International Filing Date:</b> 19 September 1997 (19.09.97)  <b>(30) Priority Data:</b> 60/026,823 27 September 1996 (27.09.96) US  <b>(71) Applicants (for all designated States except US):</b> AMERICAN CYANAMID COMPANY [US/US]; Five Giralda Farms, Madison, NJ 07940 (US). THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> UDEM, Stephen, A. [US/US]; Apartment 6F/G, 155 West 70th Street, New York, NY 10023 (US). SIDHU, Mohinderjit, S. [US/US]; 3 Aspen Court, Highland Mills, NY 10930 (US). TATEM, Joanne, M. [US/US]; 62 Douglas Drive, Towaco, NJ 07082 (US). MURPHY, Brian, R. [US/US]; 5410 Tuscarawas Road, Bethesda, MD 20816 (US). RANDOLPH, Valerie, B. [US/US]; 535 Pine Brook Road, Lincoln Park, NJ 07035 (US).	<b>(74) Agents:</b> GORDON, Alan, M. et al.; American Home Products Corporation, Patent Law Dept. - 2B, One Campus Drive, Parsippany, NJ 07054 (US).  <b>(81) Designated States:</b> AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, RO, RU, SD, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ZW, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>	
<b>(54) Title:</b> 3' GENOMIC PROMOTER REGION AND POLYMERASE GENE MUTATIONS RESPONSIBLE FOR ATTENUATION IN VIRUSES OF THE ORDER DESIGNATED MONONEGAVIRALES		
<b>(57) Abstract</b>  Isolated, recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA viruses of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene are described. Vaccines are formulated comprising such viruses and a physiologically acceptable carrier. The vaccines are used for immunizing an individual to induce protection against a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						



3' GENOMIC PROMOTER REGION AND POLYMERASE GENE  
MUTATIONS RESPONSIBLE FOR ATTENUATION IN VIRUSES  
OF THE ORDER DESIGNATED MONONEGAVIRALES

5

Field Of The Invention

This invention relates to isolated, recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA viruses of the Order designated Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene. This invention was made with Government support under a grant awarded by the Public Health Service. The Government has certain rights in the invention.

15

Background Of The Invention

20

25

30

35

Enveloped, negative-sense, single stranded RNA viruses are uniquely organized and expressed. The genomic RNA of negative-sense, single stranded viruses serves two template functions in the context of a nucleocapsid: as a template for the synthesis of messenger RNAs (mRNAs) and as a template for the synthesis of the antigenome (+) strand. Negative-sense, single stranded RNA viruses encode and package their own RNA dependent RNA Polymerase. Messenger RNAs are only synthesized once the virus has been uncoated in the infected cell. Viral replication occurs after synthesis of the mRNAs and requires the continuous synthesis of viral proteins. The newly synthesized antigenome (+) strand serves as the template for generating further copies of the (-) strand genomic RNA.

- 2 -

The polymerase complex actuates and achieves transcription and replication by engaging the cis-acting signals at the 3' end of the genome, in particular, the promoter region. Viral genes are then transcribed from the genome template unidirectionally from its 3' to its 5' end. There is always less mRNA made from the downstream genes (e.g., the polymerase gene (L)) relative to their upstream neighbors (i.e., the nucleoprotein gene (N)). Therefore, there is always a gradient of mRNA abundance according to the position of the genes relative to the 3'-end of the genome.

Based on the revised reclassification in 1993 by the International Committee on the Taxonomy of Viruses, an Order, designated Mononegavirales, has been established. This Order contains three families of enveloped viruses with single stranded, nonsegmented RNA genomes of minus polarity (negative-sense). These families are the Paramyxoviridae, Rhabdoviridae and Filoviridae. The family Paramyxoviridae has been further divided into two subfamilies, Paramyxovirinae and Pneumovirinae. The subfamily Paramyxovirinae contains three genera, Paramyxovirus, Rubulavirus and Morbillivirus. The subfamily Pneumovirinae contains the genus Pneumovirus.

The new classification is based upon morphological criteria, the organization of the viral genome, biological activities and the sequence relationships of the proteins. The morphological distinguishing feature among enveloped viruses for the subfamily Paramyxovirinae is the size and shape of the nucleocapsids (diameter 18nm, 1mm in length, pitch of 5.5 nm), which have a left-handed helical symmetry. The biological criteria are: 1) antigenic cross-reactivity between members of a genus, and 2) the presence of neuraminidase activity in the genera Paramyxovirus,

- 3 -

Rubulavirus and its absence in genus Morbillivirus. In addition, variations in the coding potential of the P gene are considered, as is the presence of an extra gene (SH) in Rubulaviruses.

5           Pneumoviruses can be distinguished from  
Paramyxovirinae morphologically because they contain  
narrow nucleocapsids. In addition, pneumoviruses have  
major differences in the number of protein-encoding  
10           cistrons (10 in pneumoviruses versus 6 in  
Paramyxovirinae) and an attachment protein (G) that is  
very different from that of Paramyxovirinae. Although  
the paramyxoviruses and pneumoviruses have six proteins  
that appear to correspond in function (N, P, M, G/H/HN,  
15           F and L), only the latter two proteins exhibit  
significant sequence relatedness between the two  
subfamilies. Several pneumoviral proteins lack  
counterparts in most of the paramyxoviruses, namely the  
nonstructural proteins NS1 and NS2, the small  
20           hydrophobic protein SH, and a second protein M2. Some  
paramyxoviral proteins, namely C and V, lack  
counterparts in pneumoviruses. However, the basic  
genomic organization of pneumoviruses and  
paramyxoviruses is the same. The same is true of  
25           rhabdoviruses and filoviruses. Table 1 presents the  
current taxonomical classification of these viruses,  
together with examples of each genus.

Table 1

30           Classification of Nonsegmented, negative-sense, single  
stranded RNA Viruses of the Order Mononegavirales  
Family Paramyxoviridae

            Subfamily Paramyxovirinae

                Genus Paramyxovirus

                    Sendai virus (mouse parainfluenza virus  
35                      type 1)

SUBSTITUTE SHEET (RULE 26)

- 4 -

Human parainfluenza virus (PIV) types 1  
and 3

Bovine parainfluenza virus (BPV) type 3

Genus *Rubulavirus*

5 Simian virus 5 (SV) (Canine  
parainfluenza virus type 2)

Mumps virus

Newcastle disease virus (NDV) (avian  
Paramyxovirus 1)

10 Human parainfluenza virus types 2, 4a  
and 4b

Genus *Morbillivirus*

Measles virus (MV)

Dolphin Morbillivirus

15 Canine distemper virus (CDV)

Peste-des-petits-ruminants virus

Phocine distemper virus

Rinderpest virus

Subfamily *Pneumovirinae*

20 Genus *Pneumovirus*

Human respiratory syncytial virus (RSV)

Bovine respiratory syncytial virus

Pneumonia virus of mice

Turkey rhinotracheitis virus

25 Family *Rhabdoviridae*

Genus *Lyssavirus*

Rabies virus

Genus *Vesiculovirus*

Vesicular stomatitis virus

30 Genus *Ephemerovirus*

Bovine ephemeral fever virus

Family *Filoviridae*

Genus *Filovirus*

Marburg virus

35

SUBSTITUTE SHEET (RULE 26)

- 5 -

For many of these viruses, no vaccines of any kind are available. Thus, there is a need to develop vaccines against such human and animal pathogens. Such vaccines would have to elicit a protective immune response in the recipient. The qualitative and quantitative features of such a favorable response are extrapolated from those seen in survivors of natural virus infection, who, in general, are protected from reinfection by the same or highly related viruses for some significant duration thereafter.

A variety of approaches can be considered in seeking to develop such vaccines, including the use of: (1) purified individual viral protein vaccines (subunit vaccines); (2) inactivated whole virus preparations; and (3) live, attenuated viruses.

Subunit vaccines have the desirable feature of being pure, definable and relatively easily produced in abundance by various means, including recombinant DNA expression methods. To date, with the notable exception of hepatitis B surface antigen, viral subunit vaccines have generally only elicited short-lived and/or inadequate immunity, particularly in naive recipients.

Formalin inactivated whole virus preparations of polio (IPV) and hepatitis A have proven safe and efficacious. In contrast, immunization with similarly inactivated whole viruses such as respiratory syncytial virus and measles virus vaccines elicited unfavorable immune responses and/or response profiles which predisposed vaccinees to exaggerated or aberrant disease when subsequently confronted with the natural or "wild-type" virus.

Early attempts (1966) to vaccinate young children using a parenterally administered formalin-inactivated RSV vaccine. Unfortunately, several field

- 6 -

trials of this vaccine revealed serious adverse reactions -- the development of a severe illness with unusual features following subsequent natural infection with RSV (Bibliography entries 1,2). It has been suggested that this formalinized RSV antigen elicited an abnormal or unbalanced immune response profile, predisposing the vaccinee to RSV disease (3,4).

Thereafter, live, attenuated RSV vaccine candidates were generated by cold passage or chemical mutagenesis. These RSV strains were found to have reduced virulence in seropositive adults. Unfortunately, they proved either over or under-attenuated when given to seronegative infants; in some cases, they also were found to lack genetic stability (5,6). Another vaccination approach using parenteral administration of live virus was ineffective and efforts along this line were discontinued (7). Notably, these live RSV vaccines were never associated with disease enhancement as observed with the formalin-inactivated RSV vaccine described above. Currently, there are no RSV vaccines approved for administration to humans, although clinical trials are now in progress with cold-passaged, chemically mutagenized strains of RSV designated A2 and B-1.

Appropriately attenuated live derivatives of wild-type viruses offer a distinct advantage as vaccine candidates. As live, replicating agents, they initiate infection in recipients during which viral gene products are expressed, processed and presented in the context of the vaccinee's specific MHC class I and II molecules, eliciting humoral and cell-mediated immune responses, as well as the coordinate cytokine patterns, which parallel the protective immune profile of survivors of natural infection.

- 7 -

This favorable immune response pattern is contrasted with the delimited responses elicited by inactivated or subunit vaccines, which typically are largely restricted to the humoral immune surveillance arm. Further, the immune response profile elicited by some formalin inactivated whole virus vaccines, e.g., measles and respiratory syncytial virus vaccines developed in the 1960's, have not only failed to provide sustained protection, but in fact have led to a predisposition to aberrant, exaggerated, and even fatal illness, when the vaccine recipient later confronted the wild-type virus.

While live, attenuated viruses have highly desirable characteristics as vaccine candidates, they have proven to be difficult to develop. The crux of the difficulty lies in the need to isolate a derivative of the wild-type virus which has lost its disease-producing potential (i.e., virulence), while retaining sufficient replication competence to infect the recipient and elicit the desired immune response profile in adequate abundance.

Historically, this delicate balance between virulence and attenuation has been achieved by serial passage of a wild-type viral isolate through different host tissues or cells under varying growth conditions (such as temperature). This process presumably favors the growth of viral variants (mutants), some of which have the favorable characteristic of attenuation. Occasionally, further attenuation is achieved through chemical mutagenesis as well.

This propagation/passage scheme typically leads to the emergence of virus derivatives which are temperature sensitive, cold-adapted and/or altered in their host range -- one or all of which are changes

from the wild-type, disease-causing viruses -- i.e., changes that may be associated with attenuation.

Several live virus vaccines, including those for the prevention of measles and mumps (which are paramyxoviruses), and for protection against polio and rubella (which are positive strand RNA viruses), have been generated by this approach and provide the mainstay of current childhood immunization regimens throughout the world.

Nevertheless, this means for generating attenuated live virus vaccine candidates is lengthy and, at best, unpredictable. The resulting mutants with desirable attenuation characteristics are selected on the basis of their ability to grow in animal cells, and even appear to be attenuated in animal models. However, all too often they remain either under- or overattenuated in the human or animal host for whom they are intended as vaccine candidates.

Even as to current vaccines in use, there is still a need for more efficacious vaccines. For example, the current measles vaccines provide reasonably good protection. However, recent measles epidemics suggest deficiencies in the efficacy of current vaccines. Despite maternal immunization, high rates of acute measles infection have occurred in children under age one, reflecting the vaccines' inability to induce anti-measles antibody levels comparable to those developed following wild-type measles infection (8,9,10). As a result, vaccine-immunized mothers are less able to provide their infants with sufficient transplacentally-derived passive antibodies to protect the newborns beyond the first few months of life.

SUBSTITUTE SHEET (RULE 26)



- 9 -

Acute measles infections in previously immunized adolescents and young adults point to an additional problem. These secondary vaccine failures indicate limitations in the current vaccines' ability to induce and maintain antiviral protection that is both abundant and long-lived (11,12,13). Recently, yet another potential problem was revealed. The hemagglutinin protein of wild-type measles isolated over the past 15 years has shown a progressively increasing distance from the vaccine strains (14). This "antigenic drift" raises legitimate concerns that the vaccine strains may not contain the ideal antigenic repertoire needed to provide optimal protection. Thus, there is a need for improved vaccines.

Rational vaccine design would be assisted by a better understanding of these viruses, in particular, by the identification of the virally encoded determinants of virulence as well as those genomic changes which are responsible for attenuation.

#### Summary Of The Invention

Accordingly, it is an object of this invention to identify those regions of the genome of the RNA viruses of the Order Mononegavirales where mutations result in attenuation of those viruses.

It is a further object of this invention to produce recombinantly-generated viruses which incorporate such attenuating mutations in their genomes.

It is still a further object of this invention to formulate vaccines containing such attenuated viruses.

These and other objects of the invention as discussed below are achieved by the generation and

- 10 -

isolation of recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA viruses of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene.

In the case of measles virus, at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 26 (A → T), nucleotide 42 (A → T or A → C) and nucleotide 96 (G → A), where these nucleotides, as well as others delineated in this application (unless stated otherwise), are presented in positive strand, antigenomic, that is, message (coding) sense, and at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 331 (isoleucine → threonine), 1409 (alanine → threonine), 1624 (threonine → alanine), 1649 (arginine → methionine), 1717 (aspartic acid → alanine), 1936 (histidine → tyrosine), 2074 (glutamine → arginine) and 2114 (arginine → lysine).

In the case of human parainfluenza virus type 3, at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 23 (T → C), nucleotide 24 (C → T), nucleotide 28 (G → T) and nucleotide 45 (T → A), and at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 942 (tyrosine → histidine), 992 (leucine →

- 11 -

phenylalanine), 1292 (leucine → phenylalanine), and 1558 (threonine → isoleucine).

In the case of human respiratory syncytial virus subgroup B, at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 4 (C → G) and the insertion of an additional A in the stretch of A's at nucleotides 6-11, and at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 353 (arginine → lysine), 451 (lysine → arginine), 1229 (aspartic acid → asparagine), 2029 (threonine → isoleucine) and 2050 (asparagine → aspartic acid).

In another embodiment of this invention, attenuated virus is used to prepare vaccines which elicit a protective immune response against the wild-type form of the virus.

In yet another embodiment of this invention, an isolated, positive strand, antigenomic message sense nucleic acid molecule (or an isolated, negative strand genomic sense nucleic acid molecule) having the complete viral nucleotide sequence (whether of wild-type virus or virus attenuated by non-recombinant means) is manipulated by introducing one or more of the attenuating mutations described in this application to generate an isolated, recombinantly-generated attenuated virus. This virus is then used to prepare vaccines which elicit a protective immune response against the wild-type form of the virus.

In still another embodiment of this invention, such a complete wild-type or vaccine viral nucleotide sequence is used: (1) to design PCR primers for use in a PCR assay to detect the presence of the

SUBSTITUTE SHEET (RULE 26)

- 12 -

corresponding virus in a sample; or (2) to design and select peptides for use in an ELISA to detect the presence of the corresponding virus in a sample.

5

### Brief Description Of The Figures

Figure 1 depicts the passage history of the Edmonston measles virus (15). The abbreviations have the following meanings: HK - human kidney; HA - human  
10 amnion; CE(am) - chick embryo; CEF - chick embryo fibroblast; DK - dog kidney; WI-38 - human diploid cells; SK - sheep kidney; \* - plaque cloning. The number following each abbreviation represents the number of passages.

15

Figure 2 depicts a map of the measles virus genome showing putative cis-acting regulatory elements at and near the genome and antigenome termini. Top - a  
schematic map of the measles virus genome, beginning at the 3' end with 52 nucleotides of leader sequence (l)  
20 and ending at the 5' terminus with 37 nucleotides of trailer sequence (t). Gene boundaries are denoted by vertical bars; below each gene is the number of cistronic nucleotides. Bottom - an expanded schematic view of the 3' extended genomic promoter regions of  
25 genome and antigenome, showing the position and sequence of the two highly conserved domains, A and B. The intervening intergenic trinucleotide is denoted as well. Nascent 5' RNAs encompassing the A' to B' regions are presumed to contain the regulatory sequence  
30 at which the N protein encapsidation initiates.

Figure 3 depicts a genetic map of the RSV subgroup B wild-type strains designated 2B and 18537 (top portion), the intergenic sequences of those strains (middle portion) and the 68 nucleotide overlap  
35 between the M2 and L genes (bottom portion). The RSV

- 13 -

2B strain has six fewer nucleotides in the G gene,  
encoding two fewer amino acid residues in the G  
protein, as compared to the 18537 strain. The 2B  
strain has 145 nucleotides in the 5' trailer region, as  
5 compared to 149 nucleotides in the 18537 strain. The  
2B strain has one more nucleotide in each of the NS-1,  
NS-2 and N genes, and one fewer nucleotide in each of  
the M and F genes, as compared to the 18537 strain.

10

#### Detailed Description Of The Invention

Transcription and replication of negative-  
sense, single stranded RNA viral genomes are achieved  
through the enzymatic activity of a multimeric protein  
15 acting on the ribonucleoprotein core (nucleocapsid).  
Naked genomic RNA cannot serve as a template. Instead,  
these genomic sequences are recognized only when they  
are entirely encapsidated by the N protein into the  
nucleocapsid structure. It is only in that context  
20 that the genomic and antigenomic terminal promoter  
sequences are recognized to initiate the  
transcriptional or replication pathways.

All paramyxoviruses require the two viral  
proteins, L and P, for these polymerase pathways to  
25 proceed. The pneumoviruses, including RSV, also  
require the transcription elongation factor, M2, for  
the transcriptional pathway to proceed efficiently.  
Additional cofactors may also play a role, including  
perhaps the virus-encoded NS1 and NS2 proteins, as well  
30 as perhaps host-cell encoded proteins.

However, considerable evidence indicates that  
it is the L protein which performs most, if not all,  
the enzymatic processes associated with transcription  
and replication, including initiation, and termination  
35 of ribonucleotide polymerization, capping and

- 14 -

polyadenylation of mRNA transcripts, methylation and perhaps specific phosphorylation of P proteins. The L protein's central role in genomic transcription and replication is supported by its large size, sensitivity to mutations, and its catalytic level of abundance in the transcriptionally active viral complex (16).

These considerations led to the proposal that L proteins consist of a linear array of domains whose concatenated structure integrates discrete functions (17). Indeed, three such delimited, discrete elements within the negative-sense virus L protein have been identified based on their relatedness to defined functional domains of other well-characterized proteins. These include: (1) a putative RNA template recognition and/or phosphodiester bond formation domain; (2) an RNA binding element; and (3) an ATP binding domain. All prior studies of L proteins of nonsegmented negative-sense, single stranded RNA viruses have revealed these putative functional elements (17).

Without being bound by the following, it is reasonable to presume that these non-protein coding, promoter and other cis-acting genomic regulatory domains are important determinants of the efficiency with which transcription and replication by measles virus (MV) and other viruses of the Order Mononegavirales are actualized, in association with the L protein, and that they may therefore be virulence determinants for these viruses as well.

In summary, the invention is believed to encompass a coordinate set of changes between the cis-acting regulatory signal (3' genomic promoter region) and the polymerase gene (L) which results in attenuation of the virus while retaining sufficient ability of the virus to replicate. Attenuation is

- 15 -

optimized by rational mutations of the 3' genomic promoter region and the polymerase gene, which provide the desired balance of replication efficiency: so that the virus vaccine is no longer able to produce disease, yet retains its capacity to infect the vaccinee's cells, to express sufficiently abundant gene products to elicit the full spectrum and profile of desirable immune responses, and to reproduce and disseminate sufficiently to maximize the abundance of the immune response elicited.

Without being bound by the following, attenuating mutations in the extended promoter (3' genomic promoter region) and in the polymerase gene are believed to affect the display of cis-acting signals and the conformation of the polymerase complex engaging these signals. For example, when encapsidated, the promoter RNA is coiled in a helical array. Changes in promoter sequence may affect the relative positions at which the conserved signals are displayed relative to one another. Specifically, the measles wild-type 3' genomic promoter region has a pyrimidine (uracil) at positions 26 and 42 (the antigenomic message sense sequences have the purine adenine). The vaccine strains have purines at those positions (the antigenomic message sense sequences have the corresponding pyrimidines; see Table 3 in Example 1 below). The larger purines may change the distance and/or angular display between the conserved domains of the promoter (e.g., in measles, positions 1-11 and 87-98), resulting in an altered spatial presentation of the cis-acting signals to the polymerase.

Animal studies have demonstrated a decrease in viral replication sufficient to avoid illness but adequate to elicit the desired immune response. This likely represents a decrease in transcription, a

- 16 -

decrease in gene expression of virally encoded proteins, a decrease in antisense templates and, therefore, the production of fewer new genomes. The resulting attenuated viruses are significantly less virulent than the wild-type.

The attenuating mutations described herein may be introduced into viral strains by two methods:

(1) Conventional means such as chemical mutagenesis during virus growth in cell cultures to which a chemical mutagen has been added, selection of virus that has been subjected to passage at suboptimal temperature in order to select temperature sensitive and/or cold adapted mutations, identification of mutant virus that produce small plaques in cell culture, and passage through heterologous hosts to select for host range mutations. These viruses are then screened for attenuation of their biological activity in an animal model. Attenuated viruses are subjected to nucleotide sequencing of their 3' genomic promoter region and polymerase genes to locate the sites of attenuating mutations. Once this has been done, method (2) is then carried out.

(2) A preferred means of introducing attenuating mutations comprises making predetermined mutations using site-directed mutagenesis. These mutations are identified either by method (1) or by reference to closely-related viruses whose attenuating mutations are already known. One or more mutations are introduced into each of the 3' genomic promoter region and the polymerase gene. Cumulative effects of different combinations of coding and non-coding changes can also be assessed.

The mutations to the 3' genomic promoter region and polymerase gene are introduced by standard recombinant DNA methods into a DNA copy of the viral



- 17 -

genome. This may be a wild-type or a modified viral genome background (such as viruses modified by method (1)), thereby generating a new virus. Infectious clones or particles containing these attenuating mutations are generated using the cDNA "rescue" system, which has been applied to a variety of viruses, including Sendai virus (18); measles virus (19); respiratory syncytial virus (20); rabies (21); vesicular stomatitis virus (VSV) (15); and rinderpest virus (23); these references are hereby incorporated by reference. See, for measles virus rescue, published International patent application WO 97/06270, designating the United States (24); for PIV-3 rescue, U.S. provisional patent application 60/047575 (25); for RSV rescue, published International patent application WO 97/12032, designating the United States (26); these applications are hereby incorporated by reference.

Briefly, all Mononegavirales rescue systems can be summarized as follows: Each requires a cloned DNA equivalent of the entire viral genome placed between a suitable DNA-dependent RNA polymerase promoter (e.g., the T7 RNA polymerase promoter) and a self-cleaving ribozyme sequence (e.g., the hepatitis delta ribozyme) which is inserted into a propagatable bacterial plasmid. This transcription vector provides the readily manipulable DNA template from which the RNA polymerase (e.g., T7 RNA polymerase) can faithfully transcribe a single-stranded RNA copy of the viral antigenome (or genome) with the precise, or nearly precise, 5' and 3' termini. The orientation of the viral genomic DNA copy and the flanking promoter and ribozyme sequences determine whether antigenome or genome RNA equivalents are transcribed. Also required for rescue of new virus progeny are the virus-specific trans-acting proteins needed to encapsidate the naked,

- 18 -

single-stranded viral antigenome or genome RNA transcripts into functional nucleocapsid templates: the viral nucleocapsid (N or NP) protein, the polymerase-associated phosphoprotein (P) and the polymerase (L) protein. These proteins comprise the active viral RNA-dependent RNA polymerase which must engage this nucleocapsid template to achieve transcription and replication.

The trans-acting proteins required for measles virus rescue are the encapsidating protein N, and the polymerase complex proteins, P and L. For PIV-3, the encapsidating protein is designated NP, and the polymerase complex proteins are also referred to as P and L. For RSV, the virus-specific trans-acting proteins include N, P and L, plus an additional protein, M2, the RSV-encoded transcription elongation factor.

Typically, these viral trans-acting proteins are generated from one or more plasmid expression vectors encoding the required proteins, although some or all of the required trans-acting proteins may be produced within mammalian cells engineered to contain and express these virus-specific genes and gene products as stable transformants.

The typical (although not necessarily exclusive) circumstances for rescue include an appropriate mammalian cell milieu in which T7 polymerase is present to drive transcription of the antigenomic (or genomic) single-stranded RNA from the viral genomic cDNA-containing transcription vector. Either cotranscriptionally or shortly thereafter, this viral antigenome (or genome) RNA transcript is encapsidated into functional templates by the nucleocapsid protein and engaged by the required polymerase components produced concurrently from co-

- 19 -

transfected expression plasmids encoding the required virus-specific trans-acting proteins. These events and processes lead to the prerequisite transcription of viral mRNAs, the replication and amplification of new genomes and, thereby, the production of novel viral progeny, i.e., rescue.

For the rescue of rabies, VSV and Sendai, T7 polymerase is provided by recombinant vaccinia virus VTF7-3. This system, however, requires that the rescued virus be separated from the vaccinia virus by physical or biochemical means or by repeated passaging in cells or tissues that are not a good host for poxvirus. For MV cDNA rescue, this requirement is avoided by creating a cell line that expresses T7 polymerase, as well as viral N and P proteins. Rescue is achieved by transfecting the genome expression vector and the L gene expression vector into the helper cell line. Advantages of the host-range mutant of the vaccinia virus, MVA-T7, which expresses the T7 RNA polymerase, but does not replicate in mammalian cells, are exploited to rescue RSV, Rinderpest virus and MV. After simultaneous expression of the necessary encapsidating proteins, synthetic full length antigenomic viral RNA are encapsidated, replicated and transcribed by viral polymerase proteins and replicated genomes are packaged into infectious virions. In addition to such antigenomes, genome analogs have now been successfully rescued for Sendai and PIV-3 (25,27).

The rescue system thus provides a composition which comprises a transcription vector comprising an isolated nucleic acid molecule encoding a genome or antigenome of a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating

- 20 -

mutation in the RNA polymerase gene, together with at least one expression vector which comprises at least one isolated nucleic acid molecule encoding the trans-acting proteins necessary for encapsidation, transcription and replication (e.g., N, P and L for measles virus; NP, P and L for PIV-3; N, P, L and M2 for RSV). Host cells are then transformed or transfected with the at least two expression vectors just described. The host cells are cultured under conditions which permit the co-expression of these vectors so as to produce the infectious attenuated virus.

The rescued infectious virus is then tested for its desired phenotype (temperature sensitivity, cold adaptation, plaque morphology, and transcription and replication attenuation), first by *in vitro* means. The mutations at the cis-acting 3' genomic promoter region are also tested using the minireplicon system where the required trans-acting encapsidation and polymerase activities are provided by wild-type or vaccine helper viruses, or by plasmids expressing the N, P and different L genes harboring gene-specific attenuating mutations (19,28).

If the attenuated phenotype of the rescued virus is present, challenge experiments are conducted with an appropriate animal model. Non-human primates provide the preferred animal model for the pathogenesis of human disease. These primates are first immunized with the attenuated, recombinantly-generated virus, then challenged with the wild-type form of the virus. Monkeys are infected by various routes, including but not limited to intranasal, intratracheal or subcutaneous routes of inoculation (29). Experimentally infected rhesus and cynomolgus macaques have also served as animal models for studies of

- 21 -

vaccine-induced protection against measles (30). Protection is measured by such criteria as disease signs and symptoms, survival, virus shedding and antibody titers. If the desired criteria are met, the attenuated, recombinantly-generated virus is considered a viable vaccine candidate for testing in humans. The "rescued" virus is considered to be "recombinantly-generated", as are the progeny and later generations of the virus, which also incorporate the attenuating mutations.

Even if a "rescued virus is underattenuated or overattenuated relative to optimum levels for vaccine use, this is information which is valuable for developing such optimum strains.

Optimally, a codon containing an attenuating point mutation may be stabilized by introducing a second or a second plus a third mutation in the codon without changing the amino acid encoded by the codon bearing only the attenuating point mutation.

Infectious virus clones containing the attenuating and stabilizing mutations are also generated using the cDNA "rescue" system described above.

Measles virus serves as a useful model for this invention, because sequence data are now available as described herein for the disease-causing wild-type virus and for the disease-preventing vaccines which have a demonstrated history of efficacy.

Measles virus was first isolated in tissue culture in 1954 (31) from an infected patient named David Edmonston. This Edmonston strain of measles became the progenitor for many live-attenuated measles vaccines including Moraten, which is the current vaccine in the United States (Attenuvax™; Merck Sharp & Dohme, West Point, PA) and was licensed in 1968 and has proven to be efficacious.

- 22 -

Aggressive immunization programs instituted in the mid to late 1960s resulted in the precipitous drop in reported measles cases from near 700,000 in 1965 to 1500 in 1983. In parallel, other vaccine strains were also developed from the Edmonston strain (see Fig. 1), Schwarz (Institut Merieux, Lyon, France), Zagreb (Zagreb, Yugoslavia) and AIK-C (Japan). These other vaccines have also proven to be efficacious and have been used extensively. An early, reactogenic, underattenuated vaccine strain (Rubeovax™: Merck Sharp & Dohme) produced measles-like illness in children and its use thus was discontinued. It, however, was further attenuated successfully to produce the Moraten vaccine strain (see Fig. 1) (32). Live measles virus vaccine provides a success story of the development of an efficacious vaccine and provides a model for understanding the molecular mechanisms of viral vaccine attenuation among nonsegmented, negative-sense, single stranded RNA viruses.

Because of its significance as a major cause of human morbidity and mortality, measles virus (MV) has been quite extensively studied. MV is a large, relatively spherical, enveloped particle composed of two compartments, a lipoprotein membrane and a ribonucleoprotein particle core, each having distinct biological functions (33). The virion envelope is a host cell-derived plasma membrane modified by three virus-specified proteins: The hemagglutinin (H; approximately 80 kilodaltons (kD)) and fusion (F<sub>1,2</sub>; approximately 60 kD) glycoproteins project on the virion surface and confer host cell attachment and entry capacities to the viral particle (16). Antibodies to H and/or F are considered protective since they neutralize the virus' ability to initiate infection (34,35,36). The matrix (M; approximately 37

- 23 -

kD) protein is the amphipathic protein lining the membrane's inner surface, which is thought to orchestrate virion morphogenesis and thus consummate virus reproduction (37). The virion core contains the 15,894 nucleotide long genomic RNA upon which template activity is conferred by its intimate association with approximately 2600 molecules of the approximately 60 kD nucleocapsid (N) protein (38,39,40). Loosely associated with this approximately one micron long helical ribonucleoprotein particle are enzymatic levels of the viral RNA dependent RNA polymerase (L; approximately 240 kD) which in concert with the polymerase cofactor (P; approximately 70 kD), and perhaps yet other virus-specified as well as host-encoded proteins, transcribes and replicates the MV genome sequences (41).

To date, the entire nucleotide sequences (only for the Edmonston B laboratory strain and the AIK-C vaccine strain), coding potential, and organization of the MV genome have been reported (33). The six virion structural proteins are encoded by six contiguous, non-overlapping genes which are arrayed as follows: 3'-N-P-M-F-H-L-5'. Two additional MV gene products of as yet uncertain function have also been identified. These two nonstructural proteins, known as C (approximately 20 kD) and V (approximately 45 kD), are both encoded by the P gene, the former by a second reading frame within the P mRNA; the latter by a cotranscriptionally edited P gene-derived mRNA which encodes a hybrid protein having the amino terminal sequences of P and a new zinc finger-like cysteine-rich carboxy terminal domain (16).

In addition to the sequences encoding the virus-specified proteins, the MV genome contains distinctive non-protein coding domains resembling those

- 24 -

directing the transcriptional and replicative pathways of related viruses (16,42). These regulatory signals lie at the 3' and 5' ends of the MV genome and in short internal regions spanning each intercistronic boundary.

5 The former encode the putative promoter and/or regulatory sequence elements directing genomic transcription, genome and antigenome encapsidation, and replication. The latter signal transcription termination and polyadenylation of each monocistronic  
10 viral mRNA and then reinitiation of transcription of the next gene. In general, the MV polymerase complex appears to respond to these signals much as the RNA-dependent RNA polymerases of other non-segmented negative strand RNA viruses (16,42,43,44).

15 Transcription initiates at or near the 3' end of the MV genome and then proceeds in a 5' direction producing monocistronic mRNAs (40,42,45). As the polymerase traverses the MV genomic template, it encounters putative stop/start signals which, in 3' to  
20 5' order, are: a semi-conserved transcription termination/polyadenylation signal (A/G U/C UA A/U NN A<sub>n</sub>, where N may be any of the four bases) at which each monocistronic RNA is completed; a non-transcribed intergenic trinucleotide punctuation mark (CUU; except  
25 at the H:L boundary where it is CGU); and a semiconserved start signal for transcription initiation of the next gene (AGG A/G NN C/A A A/G G A/U, where N may be any of the four bases) (45,46). Since some  
30 polymerase complexes fail to reinitiate, the abundance of each MV mRNA diminishes in parallel with the distance of the encoding gene from the genomic 3' end. This mRNA gradient directly corresponds to the relative abundance of each virus-specified protein. This  
35 indicates that MV protein expression is ultimately controlled at the transcriptional level (44).



- 25 -

The 3' and 5' MV genomic termini contain non-protein coding sequences with distinct parallels to the leader and trailer RNA encoding regions of VSV (42). Nucleotides 1-55 define the region between the genomic 3' terminus and the beginning of the N gene, while 37 additional nucleotides can be found between the end of the L gene and the 5' terminus of the genome. However, unlike VSV, or even the paramyxoviruses Sendai and NDV, MV does not transcribe these terminal regions into short, unmodified (+) or (-) sense leader RNAs (47,48,49). Instead, leader readthrough transcripts, including full-length polyadenylated leader:N, leader:N:P, leader:N:P:M, and of course full-length antigenome MV RNAs are transcribed (48,49). Thus, the short leader transcript, the key operational element determining the switch from transcription to replication of the VSV single-stranded, negative polarity genome (50,51,52), seems absent in MV. This leads to consideration and exploration of alternative models for this crucial reproductive event (42).

Measles virus, as well as all other Mononegavirales except the rhabdoviruses, appears to have extended its terminal regulatory domains beyond the confines of leader and trailer encoding sequences (42). For measles, these regions encompass the 107 3' genomic nucleotides (the "3' genomic promoter region", also referred to as the "extended promoter", which comprises 52 nucleotides encoding the leader region, followed by three intergenic nucleotides, and 52 nucleotides encoding the 5' untranslated region of N mRNA) and the 109 5' end nucleotides (69 encoding the 3' untranslated region of L mRNA, the intergenic trinucleotide and 37 nucleotides encoding the trailer). Within these 3' terminal approximately 100 nucleotides

- 26 -

of both the genome and antigenome are two short regions of shared nucleotide sequence: 14 of 16 nucleotides at the absolute 3' ends of the genome and antigenome are identical. Internal to those termini, an additional region of 12 nucleotides of absolute sequence identity have been located. Their position at and near the sites at which the transcription of the MV genome must initiate and replication of the antigenome must begin, suggests that these short unique sequence domains encompass an extended promoter region.

These discrete sequence elements may dictate alternative sites of transcription initiation -- the internal domain mandating transcription initiation at the N gene start site, and the 3' terminal domain directing antigenome production (42,48,53). In addition to their regulatory role as cis-acting determinants of transcription and replication, these 3' extended genomic and antigenomic promoter regions encode the nascent 5' ends of antigenome and genome RNAs, respectively. Within these nascent RNAs reside as yet unidentified signals for N protein nucleation, another key regulatory element required for nucleocapsid template formation and consequently for amplification of transcription and replication. Figure 2 schematically shows the location and sequence of these highly conserved, putative cis-acting regulatory domains.

Terminal non-protein coding regions similar in location, size and spacing are present in the genomes of other members of the genus *Paramyxoviridae*, though only 8-11 of their absolute terminal nucleotides are shared by MV (42,54). The genomic termini of the *Morbillivirus* canine distemper virus (CDV) displays a greater degree of homology with its MV relative: 73% of the nucleotides of the leader and trailer sequences

- 27 -

of these two viruses are identical, including 16 of 18 at the absolute 3' termini and 17 of 18 at their 5' ends (55). No accessory internal CDV genomic domain-sharing homology to that of the MV extended promoter has been found. However, there is a 20 nucleotide long stretch lying between CDV genomic nucleotides 85 and 104 and 15,587 and 15,606 in which 15 of the 20 nucleotides are complementary (Gene Bank accession number AF 14953). This indicates that CDV, like MV contains an additional region within its non-coding 3' genomic and antigenomic ends that may provide important cis-acting promoter and/or regulatory signals (55).

Additionally, the precise length of the 3'-leader region (55 nucleotides) is identical among several members of the Family Paramyxoviridae (MV, CDV, PIV-3, BPV-3, SV and NDV). Further evidence for the importance of these extended, non-protein coding regions comes from analyses of a large number of distinct copy-back Defective Interfering Viruses (DIs) recently cloned from subacute sclerosing panencephalitis (SSPE) brain tissue. No DI with a stem shorter than the 95 5' terminal genomic nucleotides was found. This indicates that the minimal signals needed for MV DI RNA replication and encapsidation extend well beyond the 37 nucleotide long trailer sequence to encompass the additional internal putative regulatory domain (56).

As exemplified in part by measles virus, this invention is directed to the concept that important virulence/attenuation determinants reside in viral genomic non-protein coding regulatory regions and in the transacting transcription/replication enzyme complex with which these cis-acting elements must interact. The cis-acting domains are found both at the 3' and 5' ends of the MV genome, flanking the six

- 28 -

contiguous genes encoding viral structural proteins;  
and within the MV genome as short regions encompassing  
internal intergenic boundaries. The former encode the  
putative promoter and/or regulatory sequence elements  
directing the vital processes of genomic transcription,  
genome and antigenome encapsidation, and replication.  
The latter signal transcription termination and  
polyadenylation of each monocistronic viral mRNA and  
then reinitiation of transcription of the next gene.  
The transcription/replication enzyme, RNA dependent RNA  
polymerase molecule can modulate transcription and/or  
replicative efficiency, thereby determining the  
abundance of cytopathic viral gene products and/or  
virion progeny.

Proof of the concept of this invention for  
measles virus is obtained by first determining the  
nucleotide sequences of the non-coding regulatory  
regions (3' genomic promoter region) and the coding  
regions of the L gene (with predicted amino acid  
sequences) of the progenitor Edmonston wild-type MV  
isolate, together with available measles vaccine  
strains derived from this isolate (see Figure 1).  
Independent other wild-type isolates were examined for  
comparative purposes as well.

The nucleotide sequences (in positive strand,  
antigenomic, message sense) of four wild-type and five  
vaccine measles strains, as well as the deduced amino  
acid sequences of the RNA polymerase (L protein) of  
these measles viruses, are set forth as follows with  
reference to the appropriate SEQ ID NOS. contained  
herein:

- 29 -

<u>Virus</u>	<u>Nucleotide Sequence</u>	<u>L Protein Sequence</u>
<u>Wild-Type</u>		
Edmonston	SEQ ID NO:1	SEQ ID NO:2
1977	SEQ ID NO:3	SEQ ID NO:4
5 1983	SEQ ID NO:5	SEQ ID NO:6
Montefiore	SEQ ID NO:7	SEQ ID NO:8
<u>Vaccine</u>		
Rubeovax™	SEQ ID NO:9	SEQ ID NO:10
10 Moraten	SEQ ID NO:11	SEQ ID NO:12
Zagreb	SEQ ID NO:13	SEQ ID NO:14
AIK-C	SEQ ID NO:15	SEQ ID NO:16

Each measles virus genome listed above is 15,894 nucleotides in length. Translation of the L gene starts with the codon at nucleotides 9234-9236; the translation stop codon is at nucleotides 15783-15785. The translated L protein is 2,183 amino acids long.

Note that nucleotide 2499 of 1983 wild-type measles virus is indicated as "G" in SEQ ID NO:5. In fact, the base is actually a mixture of "G" and "C". Also note that nucleotide 2143 of Rubeovax™ vaccine virus is indicated as "T" in SEQ ID NO:9. In nine clones sequenced, this base was "T" in seven and "C" in two; thus, this base can be "T" or "C".

In addition, the Schwarz vaccine virus genome is identical to that of the Moraten vaccine virus genome (SEQ ID NO:11), except that at nucleotides 4917 and 4924, Schwarz has a "C" instead of a "T".

Nucleotide differences distinguishing the 3' genomic promoter region and nucleotide and amino acid differences distinguishing the L gene and L protein sequences of the Edmonston wild-type isolate, vaccine strains and other independently isolated wild-type

- 30 -

viruses were then compared and aligned (see Tables 3-5 in Example 1 below).

As shown in Table 3, there were three mutations from the 3' genomic promoter region (in antigenomic, message sense) of the progenitor wild-type MV isolate and the derivative vaccine strains: At nucleotide position 26, from "A" to "T"; at position 42, from "A" to "C" or from "A" to "T"; and in the case of Zagreb only, at position 96, from "G" to "A". In addition, the other examined wild-type isolates differed from both the progenitor wild-type isolate and the vaccine strains at position 50 by having "A" instead of "G".

The predicted amino acid sequences of the L genes of measles vaccine strains (Rubeovax™, Moraten, Schwarz, AIK-C and Zagreb) and wild-type isolates (1977, 1983 and Montefiore), differ from the progenitor strain (Edmonston) at 49 positions in the 2183 amino acid long open reading frame (see Tables 4 and 5 in Example 1 below).

These amino acid differences can be divided into four categories:

(1) Positions where one vaccine strain differs from the progenitor, as well as from other vaccine and wild-type strains, suggesting a potential attenuation site.

(2) Specific differences between all wild-type and all vaccine sequences; these may also constitute important attenuation sites.

(3) Residues where chronologically newer wild-types differ from older wild-types; which may be attributable to genetic drift.

(4) Positions where one or more vaccine strains and/or wild-type strains have common amino acids and differ from all the other strains; these

- 31 -

changes may represent lineage-specific, potentially attenuating changes within the vaccine strains and relatedness among the wild-type isolates, respectively.

There were four category (1) changes where one vaccine differed from the other vaccines, as well as the wild-type strains. Two of these were in Moraten and Schwarz (amino acids 331 and 2114) and two were in AIK-C (1624 and 2074). These mutations are of special interest because all of these viruses are good vaccines. Thus, these positions are sites for attenuation.

Only one position, 1717, fits into category (2), with all wild-types having aspartic acid and all vaccines having alanine. Interestingly, this position is in one of two areas where the L genes of measles and canine distemper virus (which are otherwise highly homologous) do not show exceptional conservation. This difference makes it more likely that 1717 is a key position for an attenuating mutation in measles.

There were five positions, 149, 636, 720, 2017 and 2119, where both chronologically newer wild-types (1983 and Montefiore) differ from older wild-types (Edmonston and 1977), which therefore fit into category (3). These differences suggest genetic drift rather than denoting sites of attenuating mutations. Not included in this total are 16 positions where Montefiore (the 1989 isolate) differed from the rest (see Table 5). These could be either genetic drift (category (3)) or random change (category (4)). The remaining 23 positions are category (4), with one or more of the viruses differing from the consensus.

Three of these positions (1409, 1649, 1936) are potentially attenuating category (4) mutations. These are changes where two vaccine strains have a common change from the progenitor wild-type strain.

- 32 -

These changes may be connected with the vaccine lineage leading to the Rubeovax™ and Moraten vaccines (Figure 1).

Applicants have found that their AIK-C vaccine strain nucleotide sequence differs from the published sequence (33) at 21 positions, including one insertion and one deletion. Several of these differences result in coding changes including two in the L gene (at amino acids 1477 and 2008).

Thus, the additional changes accrued within the L gene sequence as the measles progenitor strain is progressively attenuated to achieve a replicative capacity optimized for live vaccine purposes appears to be constrained and delimited. Presumably, this limited tolerance in the number and location of L gene changes is imposed not only by the need to preserve the multifunctional capacities of the polymerase, but also by the preexisting 3' promoter changes with which the evolving L protein must interact to achieve transcription and replication. In other words, optimal virus attenuation requires coordinate (i.e., linked) changes in the polymerase protein and the cis-acting regulatory elements on which it acts.

The 3'-leader displays the least tolerance for change, allowing highly selected changes during the attenuation process at nucleotide position 26 (always the change of from "A" to "T"), and at position 42 (the change of from "A" to "C" or from "A" to "T") (in antigenomic, message sense). In the case of Zagreb only, there is a single further change, from "G" to "A" at position 96, which may be important when combined with Zagreb L gene-specific changes. The 3'-leader region seems to have undergone only one instance of genetic drift since 1954, with a change of "G" to "A" at position 50 (see Table 3).



- 33 -

The net change in the 3' genomic promoter region during the attenuation process is the replacement of two pyrimidines by two purines in genomic sense in all MV vaccine strains. The co-  
5 evolution of the L gene during these attenuation processes is believed to reflect selection of subtle changes favoring reproduction of the viruses in different host cells. All the vaccine strains were grown in chick embryo (CE) or chick embryo fibroblast  
10 (CEF) cells during their attenuation process (Figure 1). In addition, some vaccine strains have been exposed to unique host cells; i.e., Zagreb vaccine was grown in dog kidney cells and human diploid cells, while the AIK-C vaccine was adapted to sheep kidney  
15 cells. Moraten and Rubeovax™ were exclusively developed in CE and CEF.

Some of the lineage-specific L gene changes (position 1649 in Rubeovax™, Moraten and Schwarz vaccines and the change at position 1717 in all  
20 vaccines) represent a subset of adaptations of the L gene to the 3'-leader to modulate the transcription/replication processes for vaccine attenuation. Additionally, individual vaccine-specific changes (category (1)) may provide additional fine tune  
25 modulation of virus replication/transcription for each vaccine strain.

Based on Table 3 and the foregoing discussion, the key attenuating mutations for the MV 3' genomic promoter region are nucleotide 26 (A → T),  
30 nucleotide 42 (A → T or A → C) and nucleotide 96 (G → A) (in antigenomic, message sense).

Based on Table 4 and the foregoing discussion, the key attenuating sites for the L protein are as follows: amino acid residues 331 (isoleucine →  
35 threonine), 1409 (alanine → threonine), 1624

SUBSTITUTE SHEET (RULE 26)

- 34 -

(threonine → alanine), 1649 (arginine → methionine), 1717 (aspartic acid → alanine), 1936 (histidine → tyrosine), 2074 (glutamine → arginine) and 2114 (arginine → lysine). It is understood that the nucleotide changes responsible for these amino acid changes are not limited to those set forth in Table 4 of Example 1 below; all changes in nucleotides which result in codons which are translated into these amino acids are within the scope of this invention.

Human parainfluenza virus type 3 (HPIV-3) is another nonsegmented, negative-sense, single stranded enveloped RNA virus. HPIV-3 belongs to the Family Paramyxoviridae (see Table 1). The genome of HPIV-3 is 15,462 nucleotides long and encodes six non-overlapping protein-encoding genes (57). Five of the genes encode a single virion structural protein each, which are designated NP (corresponding to the N protein of MV), M, F, HN (hemagglutinin-neuraminidase) and L. The sixth mRNA encodes the P protein, and by an overlapping 5' proximal open reading frame (ORF) encodes the C protein, and by the RNA editing mechanism, also encodes the D protein.

Like MV, HPIV-3 consists of a 3'-nonprotein coding leader region of 55 nucleotides, but unlike measles (where it is 37 nucleotides), it has a 44 nucleotide long 5'-trailer region. The polymerase transcribes the genome in a linear, sequential, start-stop manner which is guided by transcription signals in the RNA template.

Attempts to develop a live attenuated HPIV-3 vaccine by passaging the wild-type virus JS strain through cell culture at sub-optimal temperature has produced promising results (7,57). Several "cold passage" (cp) mutants were isolated for evaluation from different passage levels of the JS strain. One such

- 35 -

mutant resulted from 45 serial passages and was designated cp45.

This virus exhibited three interesting properties: (1) cold adaptation (ca): the ability to replicate efficiently at the suboptimal temperature of 20°C; (2) temperature sensitivity (ts): inability to replicate in vitro at temperatures greater than or equal to 39°C; and (3) small plaque morphology. This mutant appeared to be a promising vaccine candidate because: (a) its ca, ts and small plaque phenotype is stable after passage in cell culture; (b) its replication is restricted in both the upper and lower respiratory tract of hamsters; and (c) it induced significant protection in hamsters against subsequent challenge with wild-type HPIV-3 (58,59).

Evaluation of this strain in the rhesus monkey showed the attenuation mutations in cp45 to be a combination of ts and non-ts mutations (60). Subsequent evaluation in chimpanzees indicated that cp45 appeared to be satisfactorily attenuated while still able to induce a high level of protection against wild-type virus challenge (61). Later preliminary clinical evaluation of cp45 in seronegative human infants and small children suggested that this candidate vaccine strain is suitably infectious and attenuated, as well as being moderately immunogenic (61).

The cp45 strain has been grown in both fetal rhesus lung (FRhL) and Vero cells as follows: The PIV-3 cp45 virus grown in FRhL cells was prepared by inoculating confluent FRhL cell monolayers in tissue culture flasks at an MOI 0.1-1.0. The infected cell cultures were fed with EMEM medium and incubated at 32°C. About seven days later, when maximal cytopathic effects (syncytia) were observed, the virus was

SUBSTITUTE SHEET (RULE 26)

- 36 -

harvested by subjecting the cultures to one freeze-thaw cycle, pooling the fluids and then storing the virus at -70 °C.

5 The PIV-3 cp45 virus grown in Vero cells was prepared by inoculating with virus a bioreactor culture of confluent monolayers of Vero cells on microcarrier beads which was continuously stirred. The infected bioreactor culture was maintained at 30°C. The virus was harvested 4-5 days later when syncytial CPE was  
10 observed. The culture fluid containing the virus was stored at -70 °C.

The nucleotide sequences (in positive strand, antigenomic, message sense) of the HPIV-3 JS wild-type strain (89) and the cp45 vaccine strain grown in FRhL and Vero cells, as well as the deduced amino acid  
15 sequences of the RNA polymerase (L protein) of these HPIV-3 viruses, are set forth as follows with reference to the appropriate SEQ ID NOS. contained herein:

20	<u>Virus</u>	<u>Nucleotide Sequence</u>	<u>L Protein Sequence</u>
	<u>Wild-Type</u>		
	JS	SEQ ID NO:17	SEQ ID NO:18
	<u>Vaccine</u>		
25	FRhL cp45	SEQ ID NO:19	SEQ ID NO:20
	Vero cp45	SEQ ID NO:21	SEQ ID NO:22

Each PIV-3 virus genome listed above is 15,462 nucleotides in length. Translation of the L  
30 gene starts with the codon at nucleotides 8646-8648; the translation stop codon is at nucleotides 15345-15347. The translated L protein is 2,233 amino acids long.

As detailed in Example 2 and Table 6 therein  
35 below, based upon the differences between the wild-type

SUBSTITUTE SHEET (RULE 26)

- 37 -

JS strain and the FRhL-grown cp 45 mutant vaccine strain, the key attenuating mutations for the HPIV-3 3' genomic promoter region are nucleotide 23 (T → C), nucleotide 24 (C → T), nucleotide 28 (G → T) and nucleotide 45 (T → A) (in antigenomic, message sense). As also detailed in Example 2 and Table 6 therein below, key attenuating sites for the L protein of HPIV-3 include the following: amino acid residues 942 (tyrosine → histidine), 992 (leucine → phenylalanine) and 1558 (threonine → isoleucine).

In addition, the Vero-grown cp45 mutant vaccine strain contains an additional mutation resulting from a coding change in the L gene at amino acid residue 1292 (leucine → phenylalanine).

It is understood that the nucleotide changes responsible for these amino acid changes are not limited to those set forth in Example 2 below; all changes in nucleotides which result in codons which are translated into these amino acids are within the scope of this invention.

Human respiratory syncytial virus (RSV) is yet another nonsegmented, negative-sense, single stranded enveloped RNA virus. RSV belongs to the Subfamily Pneumovirinae and the genus *Pneumovirus* (see Table 1).

Two major subgroups of human RSV, designated A and B, have been identified based on reactivities of the F and G surface glycoproteins with monoclonal antibodies (62). More recently, the A and B lineages of RSV strains have been confirmed by sequence analysis (63,64). Bovine, ovine, and caprine strains of this virus have also been isolated. The host specificity of the virus is most clearly associated with the G attachment protein, which is highly divergent between

the human and the bovine/ovine strains (65,66), and may be influenced, at least in part, by receptor binding. RSV is the primary cause of serious viral pneumonia and bronchiolitis in infants and young children. Serious disease, i.e., lower respiratory tract disease (LRD), is most prevalent in infants less than six months of age. It most commonly occurs in the nonimmune infant's first exposure to RSV. RSV additionally is associated with asthma and hyperreactive airways and it is a significant cause of mortality in "high risk" children with bronchopulmonary dysplasia and congenital heart disease (CHD). It is also one of the common viral respiratory infections predisposing to otitis media in children. In adults, RSV generally presents as uncomplicated upper respiratory illness; however, in the elderly it rivals influenza as a predisposing factor in the development of serious LRD, particularly bacterial bronchitis and pneumonia. Disease is always confined to the respiratory tract, except in the severely immunocompromised, where dissemination to other organs can occur. Virus is spread to others by fomites contaminated with virus-containing respiratory secretions, and infection initiates through the nasal, oral, or conjunctival mucosa.

RSV disease is seasonal and virus is usually isolated only in the winter months, e.g., from November to April in northern latitudes. The virus is ubiquitous, and over 90% of children have been infected at least once by 2 years of age. Multiple strains cocirculate. There is no direct evidence of antigenic drift (such as that seen with influenza A viruses), but sequence studies demonstrating accumulation of amino acid changes in the hypervariable regions of the G

SUBSTITUTE SHEET (RULE 26)

- 39 -

protein and SH proteins suggest that immune pressure may drive virus evolution.

In mouse and cotton rat models, both the F and G proteins of RSV elicit neutralizing antibodies and immunization with these proteins alone provides  
5 longterm protection against reinfection (67,68).

In humans, complete immunity to RSV does not develop and reinfections occur throughout life (69,70); however, there is evidence that immune factors will  
10 protect against severe disease. A decrease in severity of disease is associated with two or more prior infections and there is evidence that children infected with one of the two major RSV subgroups may be somewhat protected against reinfection with the homologous  
15 subgroup (71), observations which suggest that a live attenuated virus vaccine may provide protection sufficient to prevent serious morbidity and mortality. Infection with RSV elicits both antibody and cell mediated immunity. Serum neutralizing antibody to the  
20 F and G proteins has been associated, in some studies, with protection from LRD, although reduction in upper respiratory disease (URD) has not been demonstrated. High levels of serum antibody in infants is associated with protection against LRD, and administration of  
25 intravenous immunoglobulin with high RSV neutralizing antibody titers has been shown to protect against severe disease in high risk children (70,72,73). The role of local immunity, and nasal antibody in particular, is being investigated.

The RSV virion consists of a  
30 ribonucleoprotein core contained within a lipoprotein envelope. The virions of pneumoviruses are similar in size and shape to those of all other paramyxoviruses. When visualized by negative staining and electron  
35 microscopy, virions are irregular in shape and range in

- 40 -

diameter from 150-300 nm (74). The nucleocapsid of this virus is a symmetrical helix similar to that of other paramyxoviruses, except that the helical diameter is 12-15 nm rather than 18nm. The envelope consists of a lipid bilayer that is derived from the host membrane and contains virally coded transmembrane surface glycoproteins. The viral glycoproteins mediate attachment and penetration and are organized separately into virion spikes. All members of paramyxovirus subfamily have hemagglutinating activity, but this function is not a defining feature for pneumoviruses, being absent in RSV but present in PVM (75). Neuraminidase activity is present in members of the genera Paramyxovirus, Rubulavirus, and is absent in Morbillivirus and Pneumovirus of mice (PVM) (75).

RSV possesses two subgroups, designated A and B. The wild-type RSV (strain 2B) genome is a single strand of negative-sense RNA of 15,218 nucleotides (SEQ ID NO:23) that are transcribed into ten major subgenomic mRNAs. Each of the ten mRNAs encodes a major polypeptide chain: Three are transmembrane surface proteins (G, F and SH); three are the proteins associated with genomic RNA to form the viral nucleocapsid (N, P and L); two are nonstructural proteins (NS1 and NS2) which accumulate in the infected cells but are also present in the virion in trace amounts and may play a role in regulating transcription and replication; one is the nonglycosylated virion matrix protein (M); and the last is M2, another nonglycosylated protein recently shown to be an RSV-specified transcription elongation factor (see Figure 3). These ten viral proteins account for nearly all of the viral coding capacity.

The viral genome is encapsidated with the major nucleocapsid protein (N), and is associated with



- 41 -

the phosphoprotein (P), and the large (L) polymerase protein. These three proteins have been shown to be necessary and sufficient for directing RNA replication of cDNA encoded RSV minigenomes (76). Further studies have shown that for transcription to proceed with full processing, the M2 protein (ORF 1) is required (74). When the M2 protein is missing, truncated transcripts predominate, and rescue of the full length genome does not occur (74).

Both the M (matrix protein) and the M2 proteins are internal virion-associated proteins that are not present in the nucleocapsid structure. By analogy with other nonsegmented negative-stranded RNA viruses, the M protein is thought to render the nucleocapsid transcriptionally inactive before packaging and to mediate its association with the viral envelope. The NS1 and NS2 proteins have only been detected in very small amounts in purified virions, and at this time are considered non-structural. Their functions are uncertain, though they may be regulators of transcription and replication. Three transmembrane surface glycoproteins are present in virions: G, F, and SH. G and F (fusion) are envelope glycoproteins that are known to mediate attachment and penetration of the virus into the host cell. In addition, these glycoproteins represent major independent immunogens (77). The function of the SH protein is unknown, although a recent report has implicated its involvement in the fusion function of the virus (78).

The genomes of two wild-type RSV subgroup B strains (2B and 18537) have now been sequenced in their entirety (see SEQ ID NOS:23 and 25, discussed below). Genomic RNA is neither capped nor polyadenylated (79). In both the virion and intracellularly, genomic RNA is tightly associated with the N protein.

- 42 -

The 3' end of the genomic RNA consists of a 44-nucleotide extragenic leader region that is presumed to contain the major viral promoter (Fig. 3). The 3' genomic promoter region is followed by ten viral genes in the order 3'-NS1-NS2-N-P-M-SH-G-F-M2-L-5' (Fig. 3). The L gene is followed by a 145-149 nucleotide extragenic trailer region (see Figure 3). Each gene begins with a conserved nine-nucleotide gene start signal 3'-GGGGCAAAAU (except for the ten-nucleotide gene start signal of the L gene, which is 3'-GGGACAAAAAU; differences underlined). For each gene, transcription begins at the first nucleotide of the signal. Each gene terminates with a semi-conserved 12-14 nucleotide gene end (3'-A G U/G U/A ANNN U/A A<sub>3,5</sub>) (where N can be any of the four bases) that directs transcription termination and polyadenylation (Fig. 3). The first nine genes are non-overlapping and are separated by intergenic regions that range in size from 3 to 56 nucleotides for RSV B strains (Fig. 3). The intergenic regions do not contain any conserved motifs or any obvious features of secondary structure and have been shown to have no influence on the preceding and succeeding gene expression in a minireplicon system (Fig. 3). The last two RSV genes overlap by 68 nucleotides (Fig. 3). The gene-start signal of the L gene is located inside of, rather than after, the M2 gene. This 68 nucleotide overlap sequence encodes the last 68 nucleotides of the M2 mRNA (exclusive of the Poly-A tail), as well as the first 68 nucleotides of the L mRNA.

Ten different species of subgenomic polyadenylated mRNAs and a number of polycistronic polyadenylated read-through transcripts are the products of genomic transcription (74).

Transcriptional mapping studies using UV light mediated

- 43 -

genomic inactivation showed that RSV genes are transcribed in their 3' to 5' order from a single promoter near the 3' end (80). Thus, RSV synthesis appears to follow the single entry, sequential transcription model proposed for all Mononegavirales (16,81). According to this model, the polymerase (L) contacts genomic RNA in the nucleocapsid form at the 3' genomic promoter region and begins transcription at the first nucleotide. RSV mRNAs are co-linear copies of the genes, with no evidence of mRNA editing or splicing.

Sequence analysis of intracellular RSV mRNAs showed that synthesis of each transcript begins at the first nucleotide of the gene start signal (74). The 5' end of the mRNAs are capped with the structure m7G(5')ppp(5')Gp (where the underlined G is the first template nucleotide of the mRNA) and the mRNAs are polyadenylated at their 3' ends (82). Both of these modifications are thought to be made co-transcriptionally by the viral polymerase. Three regions of the RSV 3' genomic promoter have been found to be important as cis acting elements (83). These regions are the first ten nucleotides (presumably acting as a promoter), nucleotides 21-25, and the gene start signal located at nucleotides 45-53 (83). Unlike other Paramyxovirinae, such as measles, Sendai and PIV-3, the remainder of the leader and non-coding region of NS1 gene of RSV was found to be highly tolerant of insertions, deletions and substitutions (83).

Additionally, by saturation mutagenesis (wherein each base is replaced independently by each of the other three bases and compared for translation and replication efficiencies) within the first 12 nucleotides of the 3' genomic promoter region, a U-tract located at nucleotides 6-10 was shown to be

SUBSTITUTE SHEET (RULE 26)

- 44 -

highly inhibitory to substitutions (83). In contrast, the first five nucleotides were relatively tolerant of a number of substitutions and two of them at position four were up-regulatory mutations, resulting in a four-  
5 to 20-fold increase in RSV-CAT RNA replication and transcription. Using a bi-cistronic minireplicon system, gene-start and gene-end motifs were shown to be signals for mRNA synthesis and appear to be self contained and largely independent of the nature of  
10 adjoining sequence (84).

The L gene start signal lies 68 nucleotides upstream of the M2 gene-end signal, resulting in gene overlap (Fig. 3) (74). The presence of the M2 gene-end signal within the L gene results in a high frequency of  
15 premature termination of L gene transcripts. Full length L mRNA is much less abundant and is made when the polymerase fails to recognize the M2 gene-end motif. This results in much lower transcription of L mRNA. The gene overlap seems incompatible with a model  
20 of linear sequential transcription. It is not known whether the polymerase that exits the M2 gene jumps backward to the L gene-start signal or whether there is a second, internal promoter for L gene transcription (74). It is also possible that the L gene is  
25 accessible by a small fraction of polymerases that fail to start transcription at the M2 gene-start signal and slide down the M2 gene to the L gene-start signal.

The relative abundance of each RSV mRNA decreases with the distance of its gene from the  
30 promoter, presumably due to polymerase fall-off during sequential transcription (80). Gene overlap is a second mechanism that reduces the synthesis of full length L mRNA. Also, certain mRNAs have features that might reduce the efficiency of translation. The  
35 initiation codon for SH mRNA is in a suboptimal Kozak

- 45 -

sequence context, while the G ORF begins at the second methionyl codon in the mRNA.

RSV RNA replication is thought (74) to follow the model proposed from studies with vesicular stomatitis virus and Sendai virus (16,81). This involves a switch from the stop-start mode of mRNA synthesis to an antiterminator read-through mode. This results in synthesis of positive sense replication-intermediate (RI) RNA that is an exact complementary copy of genomic RNA. This serves in turn as the template for the synthesis of progeny genomes. The mechanism involved in the switch to the antiterminator mode is proposed to involve cotranscriptional encapsidation of the nascent RNA by N protein (16,81). RNA replication in RSV like other nonsegmented negative-strand RNA viruses is dependent on ongoing protein synthesis (85). Predicted RI RNA has been detected for the standard virus as well as RSV-CAT minigenome (74,85). RI RNA was 10-20 fold less abundant intracellularly than was the progeny genome both for the standard and the minigenome system. The nucleotide sequences (in positive strand, antigenomic, message sense) of various wild-type, vaccine and revertant RSV strains, as well as the deduced amino acid sequences of the RNA polymerase (L protein) of these RSV viruses, are set forth as follows with reference to the appropriate SEQ ID NOS. contained herein:

- 46 -

<u>Virus</u>	<u>Nucleotide Sequence</u>	<u>L Protein Sequence</u>
<u>Wild-Type</u>		
2B	SEQ ID NO:23	SEQ ID NO:24
18537	SEQ ID NO:25	SEQ ID NO:26
<u>Vaccine</u>		
2B33F	SEQ ID NO:27	SEQ ID NO:28
2B20L	SEQ ID NO:29	SEQ ID NO:30
<u>Revertant</u>		
2B33F TS(+)	SEQ ID NO:31	SEQ ID NO:32
2B20L TS(+)	SEQ ID NO:33	SEQ ID NO:34

Each RSV virus genome encodes an L protein that is 2,166 amino acids long. Genome length and other nucleotide information is as follows:

<u>Virus</u>	<u>Genome</u>		
<u>Wild-Type</u>	<u>Length</u>	<u>L Start Codon</u>	<u>L Stop Codon</u>
2B	15218	8502-8504	15000-15002
18537	15229	8509-8511	15007-15009
<u>Vaccine</u>			
2B33F	15219	8503-8505	15001-15003
2B20L	15219	8503-8505	15001-15003
<u>Revertant</u>			
2B33F TS(+)	15219	8503-8505	15001-15003
2B20L TS(+)	15219	8503-8505	15001-15003

As detailed in Example 3 (especially Tables 7 and 8) below, the key attenuating mutations for the RSV subgroup B 3' genomic promoter region are nucleotide 4 (C → G), and the insertion of an additional A in the stretch of A's at nucleotides 6-11 (in antigenomic

- 47 -

message sense). As also detailed in Example 3 below, the key potentially attenuating sites for the L protein of RSV are as follows: amino acid residues 353 (arginine → lysine), 451 (lysine → arginine), 1229 (aspartic acid → asparagine), 2029 (threonine → isoleucine) and 2050 (asparagine → aspartic acid). It is understood that the nucleotide changes responsible for these amino acid changes are not limited to those set forth in Example 3 below; all changes in nucleotides which result in codons which are translated into these amino acids are within the scope of this invention.

The attenuated viruses of this invention exhibit a substantial reduction of virulence compared to wild-type viruses which infect human and animal hosts. The extent of attenuation is such that symptoms of infection will not arise in most immunized individuals, but the virus will retain sufficient replication competence to be infectious in and elicit the desired immune response profile in the vaccinee.

The attenuated viruses of this invention may be used to formulate a vaccine. To do so, the attenuated virus is adjusted to an appropriate concentration and formulated with any suitable vaccine adjuvant, diluent or carrier. Physiologically acceptable media may be used as carriers. These include, but are not limited to: an appropriate isotonic medium, phosphate buffered saline and the like. Suitable adjuvants include, but are not limited to MPL™ (3-O-deacylated monophosphoryl lipid A; RIBI ImmunoChem Research, Inc., Hamilton, MT) and IL-12 (Genetics Institute, Cambridge, MA).

In one embodiment of this invention, the formulation including the attenuated virus is intended for use as a vaccine. The attenuated virus may be mixed

- 48 -

with cryoprotective additives or stabilizers such as proteins (e.g., albumin, gelatin), sugars (e.g., sucrose, lactose, sorbitol), amino acids (e.g., sodium glutamate), saline, or other protective agents. This mixture is maintained in a liquid state, or is then  
5 dessicated or lyophilized for transport and storage and mixed with water immediately prior to administration.

Formulations comprising the attenuated viruses of this invention are useful to immunize a  
10 human or animal subject to induce protection against infection by the wild-type counterpart of the attenuated virus. Thus, this invention further provides a method of immunizing a subject to induce protection against infection by an RNA virus of the  
15 Order Mononegavirales by administering to the subject an effective immunizing amount of a vaccine formulation incorporating an attenuated version of that virus as described hereinabove.

A sufficient amount of the vaccine in an  
20 appropriate number of doses must be administered to the subject to elicit an immune response. Persons skilled in the art will readily be able to determine such amounts and dosages. Administration may be by any conventional effective form, such as intranasally,  
25 parenterally, orally, or topically applied to any mucosal surface such as intranasal, oral, eye, vaginal or rectal surface, such as by an aerosol spray. The preferred means of administration is by intranasal administration.

30 In another embodiment of this invention, an isolated nucleic acid molecule having the complete viral nucleotide sequence of either the wild-type viruses or vaccine viruses described herein is used to generate oligonucleotide probes (from either positive  
35 strand antigenomic message sense or negative strand



- 49 -

complementary genomic sense) and to express peptides (from positive strand antigenomic message sense only), which are used to detect the presence of those wild-type virus and/or vaccine strains in samples of body fluids and tissues. The nucleotide sequences are used to design highly specific and sensitive diagnostic tests to detect the presence of the virus in a sample.

Polymerase chain reaction (PCR) primers are synthesized with sequences based on the viral wild-type or vaccine sequences described herein. The test sample is subjected to reverse transcription of RNA, followed by PCR amplification of selected cDNA regions corresponding to the nucleotide sequence described herein which have nucleotides which are distinct for a defined strain of virus. Amplified PCR products are identified on gels and their specificity confirmed by hybridization with specific nucleotide probes.

ELISA tests are used to detect the presence of antigens of the wild-type or vaccine viral strains. Peptides are designed and selected to contain one or more distinct residues based on the wild-type or vaccine sequences described herein. These peptides are then coupled to a hapten (e.g., keyhole limpet hemocyanin (KLH) and used to immunize animals (e.g., rabbits) for the production of monospecific polyclonal antibody. A selection of these polyclonal antibodies, or a combination of polyclonal and monoclonal antibodies can then be used in a "capture ELISA" to detect antigens produced by those viruses.

Samples of the Moraten measles virus vaccine strain were deposited by Applicants with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., under the provisions of the Budapest Treaty for the Deposit of Microorganisms for the Purposes of Patent Procedures

- 50 -

("Budapest Treaty") and have been assigned ATCC accession number VR2587. Samples of the HPIV-3 virus Vero-grown cp45 vaccine strain were deposited by Applicants with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., under the provisions of the Budapest Treaty and have been assigned ATCC accession number VR2588. Samples of the 2B wild-type RSV virus were deposited by Applicants with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, U.S.A., under the provisions of the Budapest Treaty and have been assigned ATCC accession number VR2586.

Given these three deposited strains and the sequence information for these and other strains provided herein, one can use site-directed mutagenesis and rescue techniques described above to introduce mutations (or restore a wild-type genotype) of all the strains described herein, as well as taking these strains and making additional mutations from the panel of mutations set forth in Tables 3, 4 and 6-8 below.

In order that this invention may be better understood, the following examples are set forth. The examples are for the purpose of illustration only and are not to be construed as limiting the scope of the invention.

#### Examples

Standard molecular biology techniques are utilized according to the protocols described in Sambrook et al. (86).

- 51 -

Example 1Measles

5 Moraten MV vaccine virus was grown once,  
directly from the Attenuvax™ vaccine vial (Lot #0716B),  
the Schwarz vaccine virus was grown once (Lot  
96G04/M179 G41D), while the Zagreb and Rubeovax™  
vaccine viruses were each grown twice in the Vero cells  
before RNAs were made for sequence analysis. MV  
10 wildtype isolate Montefiore (56) was passed 5-6 times  
in Vero cells before extraction of RNA materials and  
similarly, MV wildtype isolates 1977, 1983 (14) were  
grown 5-7 times before extracting materials for  
analysis. Edmonston wild-type isolate received from  
15 Dr. J. Beeler (CBER) (see Fig. 1) was the original  
Edmonston isolate already passaged seven times in human  
kidney cells and three times in Vero cells before  
receipt and further passaged once in Vero cells before  
using for sequence analysis.

20 RNA was prepared by infecting Vero cells at a  
multiplicity of infection (m.o.i.) of 0.1 to 1.0 and  
allowed to reach maximum cytopathology before being  
harvested. Total RNA from measles virus-infected cells  
was extracted using Trizol™ reagent (Gibco-BRL).

25 The total RNA isolated from Vero cell passage  
material was amplified by the Reverse Transcriptase-PCR  
(Perkin-Elmer/Cetus) procedure using measles (Edmonston  
B strain (19)) specific primer pairs spanning the 3'  
and 5' promoter regions and the L gene of the viral  
30 genome. Table 2 presents these primer sequences. The  
primers of SEQ ID NOS:35-54, 74, 77 and 78 are in  
antigenomic message sense. The primers of SEQ ID  
NOS:55-73, 75, 76 and 79 are in genomic negative-sense.

- 52 -

Table 2  
Primers for PCR and Sequencing MV L Genes  
and Genomic Termini

5	9047 CATATCACTCACTCTGGGATGGAG <sub>9070</sub>	(SEQ ID NO:35)
	9371 TCAGAACATCAAGCACCGCC <sub>9390</sub>	(SEQ ID NO:36)
	9741 ACAGTCAAGACTGAGATGAG <sub>9760</sub>	(SEQ ID NO:37)
	10001 AAGAGTCAGATACATGTGGA <sub>10020</sub>	(SEQ ID NO:38)
	10351 ACATGAATCAGCCTAAAGTC <sub>10370</sub>	(SEQ ID NO:39)
10	10674 CCGAAAGAGTTCCTGCGTTACGACC <sub>10698</sub>	(SEQ ID NO:40)
	11083 CAGTCCACACAAGTACCAGG <sub>11102</sub>	(SEQ ID NO:41)
	11461 GTCAGAAGCTGTGGACCATC <sub>11480</sub>	(SEQ ID NO:42)
	11841 AATATTGCTACAACAATGGC <sub>11860</sub>	(SEQ ID NO:43)
	12196 ACTCTTCATTCTTAGACTGG <sub>12215</sub>	(SEQ ID NO:44)
15	12542 GTCCAATTATGACTATGAAC <sub>12561</sub>	(SEQ ID NO:45)
	12891 AGAACAGACATGAAGCTTGC <sub>12910</sub>	(SEQ ID NO:46)
	13232 CCAACAAGGAATGCTTCTAG <sub>13251</sub>	(SEQ ID NO:47)
	13551 ACAGCACTATCTATGATTGACCTGG <sub>13575</sub>	(SEQ ID NO:48)
	13930 GCAACATGGTTTACACATGC <sub>13949</sub>	(SEQ ID NO:49)
20	14280 AGATTGAGAGTTGATCCAGG <sub>14299</sub>	(SEQ ID NO:50)
	14629 AGGAGATACTTAAACTAAGC <sub>14648</sub>	(SEQ ID NO:51)
	14981 TAAGCTTATGCCTTTCAGCG <sub>15000</sub>	(SEQ ID NO:52)
	15337 TTAACGGACCTAAGCTGTGC <sub>15356</sub>	(SEQ ID NO:53)
	15671 GAAACAGATTATTATGACGG <sub>15690</sub>	(SEQ ID NO:54)
25	9290 CGGGCTATCTAGGTGAACTTCAGG <sub>9267</sub>	(SEQ ID NO:55)
	9500 ATTTGGATATGGAATATGAG <sub>9481</sub>	(SEQ ID NO:56)
	9840 ACTCAACTGAACTACCAAGTG <sub>9821</sub>	(SEQ ID NO:57)
	10181 AAGAACATCATGTATTTTCAG <sub>10162</sub>	(SEQ ID NO:58)
30	10549 TTATCAACGCACTGCTCATG <sub>10530</sub>	(SEQ ID NO:59)
	10919 ATTTTCAGCAATCACTTGGCATGCC <sub>10895</sub>	(SEQ ID NO:60)
	11280 GCCTCTGTGCAAACAAGCTG <sub>11261</sub>	(SEQ ID NO:61)
	11638 TCTCTAGTTACTCTAGCAGC <sub>11619</sub>	(SEQ ID NO:62)
	12010 AGGTCGTTGTTTGTGAGGAG <sub>11991</sub>	(SEQ ID NO:63)
35	12361 TCGTCCTCTTCTTTACTGTGTC <sub>12342</sub>	(SEQ ID NO:64)

SUBSTITUTE SHEET (RULE 26)

- 53 -

12689 CCGTCCTCGAGCTAGCCTCG<sub>12670</sub> (SEQ ID NO:65)  
 13052 CTCCTCCAGGCTCACATTGG<sub>13033</sub> (SEQ ID NO:66)  
 13420 GGGTTGGTACATAGCTCTGC<sub>13401</sub> (SEQ ID NO:67)  
 13767 CACCCATCTGATATTTCCCTGATGG<sub>13743</sub> (SEQ ID NO:68)  
 5 14099 TGGTTGACAGTACAAATCTG<sub>14080</sub> (SEQ ID NO:69)  
 14460 CTGAAATGGGAAGATTGTGC<sub>14441</sub> (SEQ ID NO:70)  
 14820 AGCAATCTACACTGCCTACC<sub>14801</sub> (SEQ ID NO:71)  
 15180 TCACAGATGATTCAATTATC<sub>15161</sub> (SEQ ID NO:72)  
 15530 GATCCTAGATATAAGTTCTC<sub>15511</sub> (SEQ ID NO:73)  
 10  
 1 ACCAAACAAAGTTGGGTAAGG<sub>21</sub> (SEQ ID NO:74)  
 GGGGGATCC<sub>100</sub> ATCCCTAATCCTGCTCTTGTCCC<sub>78</sub> (SEQ ID NO:75)  
 200 GATTCCTCTGATGGCTCCAC<sub>181</sub> (SEQ ID NO:76)  
 15721 TAACAGTCAAGGAGACCAAAG<sub>15741</sub> (SEQ ID NO:77)  
 15 GGAAGCTT<sub>15801</sub> AACCCCTAATCCTGCCCTAGGTGG<sub>15823</sub> (SEQ ID NO:78)  
 15894 ACCAGACAAAGCTGGGAATAGA<sub>15873</sub> (SEQ ID NO:79)

20 Overlapping PCR fragments of the complete viral genome were directly sequenced without cloning to achieve the consensus sequence, by the dideoxy terminator cycle sequencing method using both strands (ABI PRISM 377 sequencer and ABI PRISM sequencing Kit). To determine the sequence at the absolute termini, a ligation procedure described previously was used (55).

25 To test this hypothesis, the nucleotide sequences were determined for the non-protein coding regulatory regions and the L gene of the progenitor Edmonston wild-type MV isolate, for the available vaccine strains derived from this isolate, as well as  
 30 for other wild-type strains. Nucleotide (in antigenomic, message sense) and amino acid differences were then compared and aligned as set forth in Tables 3-5 (differences are in italics):

- 54 -

Table 3  
Differences in MV 3' Genomic Promoter Region  
Nucleotide Sequence

<u>Virus</u>	Nucleotide number:			
	<u>26</u>	<u>42</u>	<u>50</u>	<u>96</u>
Edmonston w-t	A	A	G	G
Vaccines:				
Rubeovax™	T	C	G	G
Moraten	T	C	G	G
Schwarz	T	C	G	G
Zagreb	T	T	G	A
AIK-C	T	C	G	G
Wild-Types:				
1977	A	A	A	G
1983	A	A	A	G
Montefiore	A	A	A	G

- 55 -

Table 4  
Differences in MV L Nucleotides and Amino Acids  
Between Edmonston Wild-Type and Vaccine Strains

	<u>331</u>	<u>1409</u>	<u>1624</u>	<u>1649</u>	<u>1717</u>	<u>1887</u>	<u>1936</u>	<u>2074</u>	<u>2114</u>
Edmonston w-t	ATT	GCA	ACC	AGG	GAT	AAC	CAT	CAA	AGA
Mutation	ACT	ACA	GCC	ATG	GCT	GAC	TAT	CGA	AAA
Edmonston w-t	I	A	T	R	D	N	H	Q	R
Rubeovax™ vac.	I	A	T	M	A	D	H	Q	R
Moraten vac.	T	A	T	M	A	D	H	Q	K
Schwarz vac.	T	A	T	M	A	D	H	Q	K
Zagreb vac.	I	T	T	R	A	N	H	Q	R
AIK-C vac.	I	T	A	R	A	N	Y	R	R

- 56 -

Table 5  
Differences in MV L Nucleotides and Amino Acids  
Between Wild-Type Strains

	<u>81</u>	<u>122</u>	<u>149</u>	<u>252</u>	<u>331</u>	<u>441</u>	<u>447</u>	<u>500</u>	<u>513</u>	<u>570</u>	<u>613</u>
Edmonston w-t	GCC	GAT	GTT	ACA	ATT	AAA	AAA	GAT	GTG	AAA	TAC
Mutation	ACC	AAT	ATT	GCA	GTT	AGA	AGA	AAT	ATG	AAT	CAC
Edmonston w-t	A	D	V	T	I	K	K	D	V	K	Y
1977 w-t	A	N	V	T	V	K	K	D	M	K	Y
1983 w-t	T	D	I	T	I	K	K	N	M	N	H
Montefiore w-t	A	D	I	A	I	R	R	D	M	K	Y
	<u>618</u>	<u>621</u>	<u>623</u>	<u>626</u>	<u>628</u>	<u>632</u>	<u>636</u>	<u>637</u>	<u>641</u>	<u>645</u>	<u>650</u>
Edmonston w-t	GTC	AGT	AGG	AGA	GCA	ATA	CAA	GTA	GAC	GAT	ATG
Mutation	GCC	AAT	AAG	AAA	GAA	GTA	CAT	ATA	AAT	AAT	ATA
Edmonston w-t	V	S	R	R	A	I	Q	V	D	D	M
1977 w-t	A	N	R	R	A	I	Q	I	D	N	M
1983 w-t	V	S	K	R	A	I	H	V	D	D	M
Montefiore w-t	V	S	R	K	E	V	H	V	N	D	I



- 57 -

Table 5 (continued)  
Differences in MV L Nucleotides and Amino Acids  
Between Wild-Type Strains

	<u>652</u>	<u>720</u>	<u>723</u>	<u>794</u>	<u>914</u>	<u>970</u>	<u>1044</u>	<u>1294</u>	<u>1569</u>	<u>1705</u>	<u>1745</u>
Edmonston w-t	GCT	ATC	TAT	CGG	CGG	GCC	GGA	AGC	GTT	ATC	AAT
Mutation	ACC	GTC	TGC	TGG	CAG	TCA	AGA	ACC	ATT	GTC	AGT
Edmonston w-t	A	I	Y	R	R	A	G	S	V	I	N
1977 w-t	A	I	C	W	Q	A	G	S	V	I	N
1983 w-t	A	V	C	R	R	S	G	T	I	I	N
Montefiore w-t	T	V	C	R	R	A	R	S	V	V	S

	<u>1860</u>	<u>1865</u>	<u>1936</u>	<u>2007</u>	<u>2013</u>	<u>2017</u>	<u>2030</u>	<u>2096</u>	<u>2119</u>	<u>2165</u>
Edmonston w-t	GTA	TTC	CAT	GAC	GAT	ACT	AAT	ATA	AAG	GTC
Mutation	ATA	TAC	TAT	GGC	GGT	ATT	AGT	GTA	CGG	ATC
Edmonston w-t	V	F	H	D	D	T	N	I	K	V
1977 w-t	V	Y	H	D	D	T	N	I	K	V
1983 w-t	V	F	Y	D	G	I	N	I	R	I
Montefiore w-t	I	F	H	G	D	I	S	V	R	V

- 58 -

Example 2PIV-3

5       A comparison of sequences (in antigenomic  
message sense) of the parental wild-type JS strain of  
PIV-3 virus and the FRhL-grown and Vero-grown forms of  
the cp45 mutant are set forth in Table 6. Where a  
codon change does not result in an amino acid change,  
10       Table 6 states "none", followed by the name of the  
unchanged amino acid.

Table 6  
Sequence Comparison of Vero- and FRhL-grown cp45 & JS strains

Gene Region	Nucleotide Position	JS	FRhL cp45	Vero cp45	Codon Change	Amino Acid Change (number in L)
3' leader	23	T	C	C		
	24	C	T	T		
	28	G	T	T		
	45	T	A	A		
	62	A	T	T		
NP UTR	397	T	C	C		
NP coding	1275	T	G	G	GTC → GCC	Val → Ala
P coding	2080	T	C	C	TCT → GCT	Ser → Ala
M coding	4347	C	A	A	AAT → AAC	none/Asn
F coding	5536	C	T	T	CCC → ACC	Pro → Thr
	6329	A	G	G	AAC → AAT	none/Asn
	6419	G	A	A	ATA → GTA	Ile → Val
HN coding	6847	T	C	C	GCA → ACA	Ala → Thr
	7956	T	C	C	GGT → GGC	none/Gly
L coding	9323	T	C	C	GTT → GCT	Val → Ala
	9971	A	G	G	TAT → TAC	none/Tyr (226)
	11469	T	C	C	GAA → GAG	none/Glu (442)
	11621	G	T	T	TAC → CAC	Tyr → His (942)
	12521	A	A	T*	TTG → TTT	Leu → Phe (992)
	12581	C	T	T	TTA → TTT	Leu → Phe (1292)
	13318	C	T	T	TTC → TTT	none/Phe (1312)
# mutations			20	21	ACT → ATT	Thr → Ile (1558)

- 60 -

Sequence analysis of the parental wild-type JS strain of PIV-3 virus and the FRhL-grown cp45 mutant showed that the latter contained 20 nucleotide changes. Four changes were in the noncoding 3'-leader region at nucleotide positions 23 (T → C), 24 (C → T), 28 (G → T) and 45 (T → A) (in antigenomic, message sense). When considered in the genomic, negative sense, the change at position 28 from the smaller pyrimidine ("C") to the larger purine ("A") may change the size of the region flanked by the conserved regions of the 3' genomic promoter region, resulting in an altered spatial presentation of the cis-acting signals to the polymerase.

Nine changes were coding changes in the NP, M, F, HN and L genes. The other seven changes were non-coding or silent changes in the NP, P, F, HN and L genes or the NP untranslated region (UTR). The cp45 mutant has been demonstrated to have poor transcription activity at non-permissive temperatures due to its *ts* phenotype (87). This *ts* phenotype has now been mapped to the viral L gene (88). Because the cp45 virus has been shown to function normally with regard to mutations in the HN and F glycoproteins (87), this supports the implication that mutations in the 3'-leader and L gene contributed to the attenuating phenotype of this virus.

Thus, the four 3' leader specific changes in FRhL-grown cp45 and the three coding changes in the L gene at amino acid positions 942 (Tyr → His), 992 (Leu → Phe) and 1558 (Thr → Ile) contributed significantly to the attenuation phenotype of the candidate cp45 vaccine strain.

Furthermore, the Vero-grown cp45 mutant vaccine strain contains an additional mutation resulting from a coding change in the L gene (marked

- 61 -

with an asterisk in Table 6) at amino acid residue 1292 (leucine → phenylalanine).

5 The first two amino acid changes in the L protein (at positions 942 and 992) map to one of the highly conserved areas among all Paramyxovirus L genes. The fourth amino acid change (at position 1558) maps to the area joining two conserved blocks corresponding to the change at amino acid 1717 in the MV vaccine strains.

10 The published literature (89) sets forth only 18 changes between the antigenomic message sense sequences of the JS and FRhL-grown cp45 strains. Sixteen of these changes were found by applicants.

15 The published literature did not report four changes found by applicants: in the 3' leader at nucleotide 45 (T → A), in the NP UTR at nucleotide 62 (A → T), or the changes in amino acids in the NP protein resulting from the changes at nucleotide 397 (T → C), leading to the amino acid change (Val → Ala) and  
20 nucleotide 1275 (T → G), leading to the amino acid change (Ser → Ala) (nucleotide changes in antigenomic, message sense). Nor did the published literature report the additional potentially attenuating mutation in the L protein found by applicants in the Vero-grown  
25 cp45 strain resulting from the change at nucleotide 12521 (A → T), leading to the change in amino acid 1292 (Leu → Phe).

- 62 -

Example 3  
RSV Subgroup B

5       The temperature-sensitive (*ts*) phenotype is  
strongly associated with attenuation *in vivo*; in  
addition, some non-*ts* mutations may also be  
attenuating. Identification of *ts* and non-*ts*  
attenuating mutations was achieved by sequence analysis  
and evaluation of *ts*, cold-adapted (*ca*), and *in vivo*  
10      growth phenotypes of RSV mutants and revertants.

      The genomes of the following five RSV 2B  
strains have now been completely sequenced: 2B parent,  
2B33F, one revertant designated 2B33F TS(+), 2B20L and  
one revertant designated 2B20L TS(+). The 2B33F and  
15      2B20L strains are *ts* and *ca* and are described in U.S.  
Serial No. 08/059,444 (90), which is hereby  
incorporated by reference. After identifying regions  
where mutations in 2B33F and 2B20L are located, nine  
additional isolates of 2B33F "revertants" obtained  
20      following *in vitro* passaging at 39°C and *in vivo*  
passaging in African Green Monkeys or chimpanzees, and  
nine additional isolates of 2B20L "revertants" obtained  
following *in vitro* passaging at 39°C have been  
sequenced in those regions. The *ts*, *ca*, and  
25      attenuation phenotypes of many of these revertants have  
now been characterized and assessed. Correlations  
between phenotype *ts*, vaccine attenuation and sequence  
changes have been identified.

      A summary of results is presented in Tables  
30      7-12.

- 63 -

Table 7  
Sequence comparison between RSV 2B and 2B33F strains

Gene/ region	Nucl. pos.†	Nucleotide changes			Amino acid changes
		RSV 2B	RSV 2B33F	RSV 2B33F TS(+), 5a revertant	
Genomic Promoter	4 6	C -	G extra A	G extra A	non-coding non-coding
M	4175 4199	T T	C C	C C	non-coding non-coding
SH	4329 4409 4420 4442 4454 4484 4497 4505 4525 4526 4542 4561 4575 4598	T T T T T T T T T T T T T T	C C C C C C C C C C C C C C	C C C C C C C C C C C C C C	Phe-Leu (10) none Ile (36) Ile-Thr (40) none His (47) none Cys (51) none Tyr (61) Stop-Gln (66) none Ser (68) Ile-Thr (75) Ile-Thr (75) Stop-Gln (81) Leu-Pro (87) Trp-Arg (92) none Thr (99)
L	9559 9853* 12186 14587 15071	G A G C A	A G A T G	A A A T G	Arg-Lys (353) Lys-Arg (451)* Asp-Asn (1229) Thr-Ile (2029) non-coding

† For 2B33F and 2B33F TS(+), nucl. pos. numbers are one larger than for 2B for M, SH & L genes

\* At pos. 9853, the Lys-Arg change has reverted back to Lys in the 2B33F TS(+) strain

- 64 -

Table 8  
Sequence comparison between RSV 2B and 2B20L strains

Gene/ region	Nucl. pos.†	Nucleotide changes			Amino acid changes
		RSV 2B	RSV 2B20L	RSV 2B20L TS(+), R1 revertant	
Genomic Promoter	4	C	G	G	non-coding*
	6	-	extra A	extra A	non-coding*
L	8963	C	T	T	none Thr (154)
	13347	A	A	G	Asn-Asp (1616)
	14587	C	T	T	Thr-Ile(2029)*
	14649	A	G	G	Asn-Asp (2050)
	14650	A	A	T	Asn-Asp-Val (2050)**

† For 2B20L and 2B20L TS(+), nucl. pos. numbers are one larger than for 2B for L gene

\* Mutation is common in 2B33F and 2B20L strains

\*\* At pos. 14650, the mutation suppresses the ts phenotype in 2B20L TS(+) revertant



Table 9  
RSV 2B, ts and Revertant Strains

Sample	Source	In Vitro Phenotype		In Vivo Growth*			
		ts	ca	Cotton Rat	AGM		
		39/32°C EOP plaque morph	20/32°C Yield	Nasal turbinate	Lungs	Nasal Wash	Bronchial Lavage
RSV 2B	Wild-type Parent Strain	0.7 (WT)	0.0001	5.5 <sup>a</sup> 3.9 <sup>b</sup> (4/4)	5.8 <sup>a</sup> 5.2 <sup>b</sup> (4/4)	5.8 <sup>a</sup> (4/4)	4.7 <sup>a</sup> (4/4)
RSV 2B33F	ca, ts mutant isolated from 2B cold-passaged x 33	0.00007 (sp/int/wt)	0.04	≤1.6 <sup>a</sup> <1.9 <sup>b</sup> (1/4)	<1.5 <sup>a</sup> <1.2 <sup>b</sup> (0/4)	3.0 <sup>a</sup> (4/4)	<0.9 <sup>a</sup> (0/4)
RSV 2B33F - 5a TS(+)	2B33F spinner passage, plaque picked at 39°C	0.5 (WT)	0.03	≤1.7 <sup>a</sup> (1/4)	3.5 <sup>a</sup> (4/4)	4.2 <sup>a</sup> (4/4)	4.0 <sup>a</sup> (4/4)
RSV 2B33F - 4a TS(+)	2B33F spinner passage, plaque picked at 39°C	0.7 (WT)	0.01	≤1.7 <sup>a</sup> (3/4)	3.8 <sup>a</sup> (4/4)	ND	ND
RSV 2B33F - 3b TS(+)	2B33F spinner passage, plaque picked at 39°C	0.5 (WT)	0.04	≤2.5 <sup>a</sup> (3/4)	2.9 <sup>a</sup> (4/4)	ND	ND
AGM pp2	2B33F-infected AGM #A2.d7 nasal wash plaque picked at 32°C	0.3 (sp,int)	0.00002	≤2.0 <sup>b</sup> (1/4)	1.6 <sup>b</sup> (4/4)	ND	ND

Table 9 (continued)  
RSV 2B, ts and Revertant Strains

Sample	Source	In Vitro Phenotype ts			In Vivo Growth* Cotton Rat				AGM
		39/32°C EOP plaque morph	20/32°C yield	ca	Nasal turbinate	Lungs	Nasal Wash	Bronchial Lavage	
AGM pp4	2B33F-infected AGM #A2, d7 nasal wash plaque picked at 32°C	0.1 (sp, int)	0.008		<1.6 <sup>b</sup> (0/4)	1.2 <sup>b</sup> (4/4)	ND	ND	ND
AGM pp6	2B33F-infected AGM #A4, d12 nasal wash plaque picked at 32°C	0.000004 (wt)	≤0.00005		≤1.5 <sup>b</sup> (1/4)	<1.1 <sup>b</sup> (0/4)	ND	ND	ND
AGM pp7	2B33F-infected AGM #A4, d12 nasal wash plaque picked at 32°C	0.000004 (sp/int/wt)	0.007		≤1.4 <sup>b</sup> (1/4)	<1.0 <sup>b</sup> (0/4)	ND	ND	ND
Chimp pp1A	2B33F-infected Chimp #1552, d4 tracheal lavage plaque picked at 32°C	0.5 (WT)	ND		ND	ND	ND	ND	ND
Chimp pp3A	2B33F-infected Chimp #1560, d6 tracheal lavage plaque picked at 32°C	0.7 (WT)	ND		2.4 <sup>c</sup> (4/4)	≤3.0 <sup>c</sup> (3/4)	ND	ND	ND
Chimp pp5A	2B33F-infected Chimp #1563, d10 nasal swab plaque picked at 32°C	0.7 (WT)	ND		≤2.3 <sup>c</sup> (3/4)	3.0 <sup>c</sup> (4/4)	ND	ND	ND

Table 9 (continued)  
RSV 2B, ts and Revertant Strains

Sample	Source	In Vitro Phenotype		In Vivo Growth*			
		ts	ca	Cotton Rat	AGM		
		39/32°C GOP plaque morph	20/32°C Yield	Nasal turbinate	Lungs	Nasal Wash	Bronchial Lavage
RSV 2B20L	ca, ts mutant isolated from 2B cold-passaged x 20	0.0002 (int/wt)	0.02	<1.9 <sup>d</sup> (0/4)	<1.3 <sup>d</sup> (0/4)	<0.7 <sup>c</sup> (0/2)	<0.7 <sup>c</sup> (0/2)
RSV 2B20L R1 TS(+)	2B20L spinner passage, plaque picked at 39°C	0.6 (WT)	ND	2.3 <sup>c</sup> (4/4)	3.5 <sup>c</sup> (4/4)	ND	ND
RSV 2B20L R2 TS(+)	2B20L spinner passage, plaque picked at 39°C	0.6 (WT)	ND	≤2.5 <sup>c</sup> (3/4)	2.7 <sup>c</sup> (4/4)	ND	ND
RSV 2B20L R9 TS(+)	2B20L spinner passage, plaque picked at 39°C	0.8 (WT)	ND	≤2.2 <sup>c</sup> (3/4)	4.0 <sup>c</sup> (4/4)	ND	ND
RSV 2B20L R10 TS(+)	2B20L spinner passage, plaque picked at 39°C	0.7 (WT)	ND	2.6 <sup>c</sup> (4/4)	3.2 <sup>c</sup> (4/4)	ND	ND

\* In Vivo growth measured in log<sub>10</sub> mean virus titer (# infected/# total)

ND = not done WT = wild-type plaque size sp = small plaque size int = intermediate plaque size

<sup>a</sup> Dose = 10<sup>6.7</sup> PFU IN <sup>b</sup> Dose = 10<sup>5.6</sup> PFU IN <sup>c</sup> Dose = 10<sup>6.3</sup> PFU IN

<sup>d</sup> Dose = 10<sup>5.3</sup> PFU IN <sup>e</sup> Dose = 10<sup>6.6</sup> PFU IN+IT <sup>f</sup> Dose = 10<sup>6.0</sup> PFU IN+IT

- 68 -

Table 10  
2B33F Revertants

	<i>ts</i> (+) <i>In vitro</i>			AGM				Chimp		
	5a	4a	3b	pp2	pp4	pp6	pp7	1A	3A	5A
base no.†										
<b>M</b>										
4176,4200	S	S	S	S	S	S	S	S	S	S
<b>SH</b>										
14 bases*	S	S	S	S	S	S	S	S	S	S
<b>L</b>										
9560	S	S	S	S	S	S	S	S	S	S
9854	2B	2B	2B	2B	S	S	S	ND	2B	2B
12187	S	S	S	S	S	S	S	S	S	S
14588	S	S	S	S	S	S	S	ND	S	S
15072	S	S	S	S	S	S	S	S	S	S
Phenotype										
<i>ts</i>	2B	2B	2B	r	r	S	S	2B	2B	2B
<i>ca</i>	S	S	S	2B	S	2B	S	ND	ND	ND
Attenuated	r	r	r	(r)	(r)	S	S	ND	r	r

† These 2B33F revertant base nos. are one larger than for 2B for M, SH and L genes

\* bases 4330,4410,4421,4443,4455,4485,4498,4506,4526,4527,4543, 4562,4576,4599

S = same base as 2B33F

2B = reversion to 2B base or complete reversion in phenotype

r = moderate reversion in phenotype

(r) = slight reversion in phenotype

ND = not done

- 69 -

Table 11  
2B20L Revertants

	TS(+) In vitro Isolates									
base no.†	R1	R2	R3A	R4A	R5A	R6A	R7A	R8A	R9A	R10A
<u>L</u>										
8964	S	S	S	S	S	S	S	S	S	S
13348	C*	S	ND	S	S	ND	S	S	S	S
14588	S	S	S	S	S	S	S	S	S	S
14650	S	S	2B	S	2B	2B	S	S	2B	2B
14651	A*	A*	S	A*	S	S	A*	A*	S	S
Phenotype										
ts	2B	2B	ND	ND	ND	ND	ND	ND	2B	2B
Attenuated	r	r	ND	ND	ND	ND	ND	ND	r	r

† These 2B20L revertant base nos. are one larger than for 2B for L genes

S = same base as 2B20L

2B = reversion to 2B base

r = moderate reversion in phenotype

\* = base change, different from 2B or 2B20L

ND = not done

- 70 -

Table 12  
RSV 2B, *ts* and Revertant Strains: Phenotype Summary

Virus Isolate	Source	In Vitro Phenotype		In Vivo Attenuation	
		<i>ts</i>	<i>ca</i>	Cotton Rat	AGM
RSV 2B	Wild-type Parent Strain	-	-	-	-
RSV 2B33F	<i>ca</i> , <i>ts</i> mutant isolated from 2B, cold-passaged x 33	++++	++	++++	+++
RSV 2B33F - 5a TS(+)	2B33F spinner passage plaque picked at 39°C	-	++	++	+
RSV 2B33F - 4a TS(+)	2B33F spinner passage plaque picked at 39°C	-	++	++	ND
RSV 2B33F - 3b TS(+)	2B33F spinner passage plaque picked at 39°C	-	++	++	ND
AGM pp2	2B33F-infected AGM A2, d7 nasal wash plaque picked at 32°C	+	-	+++	ND
AGM pp4	2B33F-infected AGM A2, d7 nasal wash plaque picked at 32°C	+	++	+++	ND
AGM pp6	2B33F-infected AGM A4, d12 nasal wash plaque picked at 32°C	++++	-	++++	ND
AGM pp7	2B33F-infected AGM A4, d12 nasal wash plaque picked at 32°C	++++	++	++++	ND
Chimp pp1A	2B33F-infected chimp #1552, d4 tracheal lavage, plaque picked at 32°C	-	ND	ND	ND
Chimp pp3A	2B33F-infected chimp #1560, d6 tracheal lavage, plaque picked at 32°C	-	ND	++	ND
Chimp pp5A	2B33F-infected chimp #1563, d10 tracheal lavage, plaque picked at 32°C	-	ND	++	ND

- 71 -

Table 12 (continued)  
 RSV 2B, *ts* and Revertant Strains: Phenotype Summary

Virus Isolate	Source	In Vitro Phenotype		In Vivo Attenuation	
		<i>ts</i>	<i>ca</i>	Cotton Rat	AGM
RSV 2B20L	<i>ca</i> , <i>ts</i> mutant isolated from 2B, cold-passaged x 20	++++	++	++++	++++
RSV 2B20L R1 TS(+)	2B20L spinner passage plaque picked at 39°C	-	ND	++	ND
RSV 2B20L R2 TS(+)	2B20L spinner passage plaque picked at 39°C	-	ND	++	ND
RSV 2B20L R9 TS(+)	2B20L spinner passage plaque picked at 39°C	-	ND	++	ND
RSV 2B20L R10 TS(+)	2B20L spinner passage plaque picked at 39°C	-	ND	++	ND

ND = not done

- = wild-type phenotype, i.e., not temperature sensitive, not cold adapted, not attenuated

+ to ++++ = increasing levels of temperature sensitivity, cold-adaptation or attenuation

- 72 -

Several significant observations can be drawn from these data:

5 a. As shown in Tables 7 (for 2B33F) and 8 (for 2B20L), there are relatively few sequence changes identified in the two mutant strains: RSV 2B33F differs from parental RSV 2B by two changes at the 3' genomic promoter region, two changes at the non-coding 5'-end of the M gene, and four coding changes plus one non-coding (poly(A) motif) change in the RNA dependent RNA polymerase coding L gene. In addition, 14 changes mapped to the SH gene alone. RSV 2B20L differs from its RSV 2B parent only at seven nucleotide positions, of which three are common with 2B33F virus, including two changes at the 3' genomic promoter and one coding change in the L gene. Two additional unique changes of 2B20L virus mapped to the coding region of the L gene. Potentially attenuating mutations at the non-coding 3' genomic promoter region and the RNA dependent RNA polymerase gene have been identified.

b. Two *ts* mutations can be identified in the L gene of the attenuated virus strains 2B33F and 2B20L:

25 (i) In 2B33F, a mutation at nucleotide position 9853 (A → G) leading to a coding change in L protein at amino acid 451 (Lys → Arg) is clearly associated with the *ts* and attenuation phenotypes. Reversion at this site alone in the 2B33F TS(+) 5a strain is responsible for complete restoration of growth at 39°C (Table 9) and partial reversion in attenuation in animals. This association with the *ts* and attenuation phenotypes was also supported by partial sequence analyses of six additional "full TS revertants" (designated 4a, 3b, pp2, 3A, 5a, 5A) isolated from cell



- 73 -

culture and from chimps, in which only the nucleotide 9853 mutation reverted (Tables 10-12) (note that one AGM (African Green Monkey) isolate which reverted at 9853 only partially reverted in *ts* phenotype). This amino acid 451 mutation (Lys → Arg) is amenable to stabilization in cDNA infectious clone constructs, by inserting a second mutation to stabilize the codon, thereby lessening the likelihood that it will revert back to Lys.

(ii) In 2B20L, a mutation at base 14,649 (A → G) leading to a coding change in the L protein (amino acid position 2,050, Asn → Asp) appears to be associated with the *ts* and attenuation phenotypes. This aspartic acid at the amino acid 2050 invariably reverts back (Asp → Asn) in TS(+) revertants or changes to a different amino acid (Asp → Val) by nucleotide substitution at position 14,650 (A → T) (Tables 8, 11). The above observation is based on complete sequence analysis on the TS(+) revertant R1 and partial sequence of several additional TS(+) revertants (R2, R4A, R7A, R8A) at selected regions (Table 11). An additional mutation is seen in the R1 revertant at nucleotide position 13,347 (amino acid 1616, Asn → Asp) associated with the above reversion. However, the effect of this mutation on the *ts* phenotype is not known; the L gene of other revertants has not been sequenced completely.

c. Three base changes are common to 2B33F and 2B20L strains of virus:

(i) A change at position 14,587 (C → T) with a corresponding change (Thr → Ile) at amino acid 2029 is

- 74 -

present in both 2B33F and 2B20L (Tables 7,8). This nucleotide "T" substitution was found to be present in 10% of the population of the progenitor RSV2B strain and may have been preferred during the attenuation process. No wildtype base "C" was found in the 2B33F and 2B20L virus.

(ii) Two mutations are seen in the 2B33F and 2B20L 3' genomic promoter region: nucleotide 4 (C → G) and the insertion of an extra A in the stretch of A's at positions 6-11 (in antigenomic, message sense). When the sequences of selected TS(+) revertants were analyzed, these mutations were seen to have been retained in the 2B33F TS(+)5a (Table 7) and the 2B20L TS(+)R1 (Table 8) revertants. These non-coding, cis-acting mutations remained associated with partial viral attenuation.

Expression using the minireplicon RSV-CAT system for the analysis of these cis-acting changes has shown the 3' genomic promoter nucleotide 4 (C → G) change to be an upregulation of transcription/replication in this in vitro system when the 2B progenitor virus or either of the 2B33F or 2B33F TS(+) provided helper L gene functions (the N, P and M2 genes are identical in these viruses).

Complementation analysis of the 2B33F 3' genomic promoter and the helper functions provided by the progenitor RSV2B virus or the 2B33F and 2B33F TS(+) viruses by this RSV-CAT minireplicon system has also been conducted. All three viruses supported both the 2B and 2B33F 3' genomic promoter mediated transcription/replication functions. However, the 2B33F and 2B33F TS(+) viruses preferred their 2B33F 3' genomic promoters. This analysis clearly shows co-evolution of 3' genomic promoter changes during the

- 75 -

vaccine attenuation process, along with the RNA dependent RNA polymerase gene. Reversion of ts phenotype in the 2B33F mutant 5a by reversion of the single L protein amino acid 451 (Arg → Lys) by  
5 sequence analysis was clearly demonstrated by support of transcription/replication functions of RSV-CAT minireplicon at 37°C. The 2B33F virus did not provide helper functions to the RSV-CAT minireplicon (with 2B or 2B33F 3' genomic promoters) at 37°C.

10 d. A biased hypermutation of SH seen in 2B33F is present in all 2B33F revertants, regardless of phenotype, and is not seen in 2B20L, which is ts, ca, and attenuated. Thus, there are no data at this time  
15 that associate this mutation with any biological phenotype.

Another wild-type RSV designated 18537 was also sequenced and compared to the sequence of the wild-type RSV 2B strain. With one exception, at all  
20 the critical residues described above, the two wild-type strains were identical. For 2B, the codon ACA at nucleotides 14586-14588 encodes a Thr at amino acid 2029 of the L protein, while for 18537, the codon ATT at nucleotides 14593-14595 encodes an Ile at amino acid  
25 2029 (the L gene start codon is at nucleotides 8509-8511 in 18537, compared to 8502-8504 in 2B).

#### Example 4

##### PCR Assay to Detect Measles Virus

30 A 21 year old patient was admitted to a hospital with a three week history of progressive non-productive cough, shortness of breath, and fever. His symptoms failed to improve following treatment with  
35 clarithromycin for seven days or after a similar course

- 76 -

of treatment with atovaquone. Concomitant complaints of right upper quadrant abdominal pain proved recalcitrant to omeprazole and antacids. Relevant past medical history included Factor VIII deficiency and HIV infection diagnosed 3-4 years prior to this hospital admission. One year earlier, he had received a booster immunization of measles-mumps-rubella (MMR) vaccine as required for college enrollment.

Bronchoalveolar lavage and transbronchial biopsies performed two days after admission to the hospital demonstrated reactive hyperplasia and alveolar lining cell desquamation with minimal chronic inflammation. No microorganisms were revealed by Gram, methenamine silver, or PAS stains. CT scans of the chest showed multiple, ill-defined, confluent nodules at the left lung base. Despite administration of empiric antimicrobials for opportunistic bacterial, mycobacterial, and fungal pathogens commonly responsible for pulmonary complications of advanced HIV disease, the patient became and remained febrile to 39°C. A left-sided pleural effusion developed; diagnostic thoracentesis showed it to be exudative but otherwise non-diagnostic. Bronchoalveolar lavage performed three weeks later only demonstrated alveolar histiocytes, some of which were hemosiderin laden, a few lymphocytes, and neutrophils. FITE, AFB, and methanamine silver stains again were negative.

Two weeks thereafter, a wedge resection of the left lung was performed through CT-guided minithoracotomy. Multiple tissue sections revealed nodular areas of acute and chronic inflammation with regions of necrosis and fibrosis. Numerous multinuclated giant cells were present, some of which contained both intracytoplasmic and intranuclear inclusions suggestive of measles virus giant cell

- 77 -

pneumonia. Special stains for bacteria, fungi, *P. carinii*, and acid fast organisms again gave negative results. Electron microscopic examination of sections of this lung biopsy revealed particles morphologically consistent with paramyxoviruses such as measles virus. Serum anti-measles IgM titers determined by a solid phase hemadsorbant assay were negative, as was a subsequent IgM capture immunoassay.

Two weeks later, Rhesus monkey kidney (RMK) tissue culture cells inoculated with the patient's lung biopsy material revealed cytopathic changes characteristic of measles virus infection. Confirmation was obtained using an immunofluorescence assay with monoclonal antibodies directed to measles virus. Based upon this diagnosis, oral ribavirin 1000mg B.I.D. was given for 14 days. Unfortunately, the patient progressively deteriorated, eventually dying two months later.

In order to ascertain the nature of the measles virus present in the patient, reverse transcription and PCR amplification of virus obtained from infected tissues were performed, followed by sequence analysis. The measles virus isolated from Rhesus monkey kidney cells inoculated with tissue from this patient's lung biopsy was propagated by two serial passages in the continuous Vero (monkey kidney) tissue culture cell line. Total infected cell RNA was extracted at the second Vero cell passage using TRIzol reagent (Life Technologies, Grand Island, NY) according to the manufacturer's protocol. Total RNA was similarly extracted from the patient's lung biopsy material. The measles virus vaccine strain (Moraten) currently used in the United States as a component of the trivalent MMR vaccines, was obtained in its univalent form (Attenuvax™, Merck, Sharpe, & Dohme).

- 78 -

This virus was passaged once in Vero cells and total vaccine infected cellular RNA then was extracted as described above.

Each of these RNA preparations was reverse  
5 transcribed (RT) to cDNA using random hexameric primers  
and Maloney murine leukemia virus reverse transcriptase  
(Perkin-Elmer/Cetus RT-PCR kit reagents, Perkin-Elmer-  
Cetus, Branchburg, NJ). The cDNA then was amplified by  
10 PCR using measles virus-specific oligodeoxynucleotide  
primer pairs whose design was based on the Edmonston  
measles virus sequence described above. These PCR  
products comprised a set of overlapping DNA fragments  
spanning the entire 15,894 nucleotide long measles  
15 genome. A consensus genomic sequence was established  
by direct analysis of each PCR product, without  
cloning, using the dideoxy terminator cycle-sequencing  
method established by the manufacturer (ABI PRISM 377  
sequencer and ABI PRISM DNA sequencing kit; Perkin-  
20 Elmer/Cetus, Foster City, CA). Both strands of the  
PCR-amplified DNA products were analyzed to eliminate  
possible sequencing ambiguities.

The nucleotide sequences of selected regions  
of the measles virus genomes present in the patient's  
viral isolate, as well as in the diseased lung tissue,  
25 were compared with that of the Moraten vaccine virus,  
as well as with the nucleotide sequences of other  
measles virus wild-type and vaccine strains. This  
sequence analysis revealed identity to the Moraten  
vaccine strain rather than demonstrating relatedness to  
30 past or currently circulating wild-type viruses or  
other measles vaccine strains.

- 79 -

Example 5ELISA to Detect RSV

5       An ELISA test is used to detect the presence  
of RSV. Peptides are designed and selected based on  
homologies to the RSV sequences described herein to be  
specific for all subgroup B strains, or for individual  
wild-type, vaccine or revertant RSV subgroup B strains  
described herein. These peptides are then coupled to  
10       KLH and used to immunize rabbits for the production of  
monospecific polyclonal antibody. A selection of these  
polyclonal antibodies, or a combination of polyclonal  
and monoclonal antibodies is then used in a "capture  
ELISA" to detect the presence of an RSV antigen.

- 80 -

Bibliography

1. Kapikian, A.Z., et al., Am. J. Epidemiol., 89, 405-421 (1969).
2. Chin, J., et al., Am. J. Epidemiol., 89, 449-463 (1969).
3. Fulginiti, V.A., et al., Am. J. Epidemiol., 89, 435-448 (1969).
4. Prince, G.A., et al., J. Virology, 57, 721-728 (1986).
5. Kim, H.W., et al., Pediatrics, 52, 56-63 (1973).
6. Hodes, D.S., et al., Proc. Soc. Exp. Biol. Med., 145, 1158-1164 (1974).
7. Belshe, R.B., and Hissom, F.K., J. Med. Virol., 10, 235-242 (1982).
8. Black, F.L., et al., Am. J. Epidemiol., 124, 442-452 (1986).
9. Lennon, J.L., and Black, F.L., J. Pediatrics, 108, 671-676 (1986).
10. Pabst, H.F., et al., Pediatr. Infect. Dis. J., 11, 525-529 (1992).
11. Centers for Disease Control, MMWR, 40, 369-372 (1991).
12. Centers for Disease Control, MMWR, 41:S6, 1-12 (1992).
13. King, G.E., et al., Pediatr. Infect. Dis. J., 10, 883-887 (1991).
14. Rota, J.S., et al., Virology, 188, 135-142 (1992).
15. Rota, J.S., et al., Virus Res., 31, 317-330 (1994).
16. Lamb, R.A., and Kolakosky, D., pages 1177-1204 of Volume 1, Fields Virology, B.N. Fields, et al., Eds. (3rd ed., Raven Press, 1996).



- 81 -

17. Sidhu, M.S., et al., Virology, 193, 50-65 (1993).
18. Garcin, D., et al., EMBO J., 14, 6087-6094 (1995).
19. Radecke, F., et al., EMBO J., 14, 5773-5783 (1995).
20. Collins, P.L., et al., Proc. Natl. Acad. Sci., USA, 92, 11563-11567 (1995).
21. Published European Patent Application No. 702,085.
22. Published International Application No. WO 96/10400.
23. Baron, M.D., and Barrett, T., J. Virology, 71, 1265-1271 (1997).
24. Published International Application No. WO 97/06270.
25. U.S. Provisional Patent Application No. 60/047575.
26. Published International Application No. WO 97/12032.
27. Kato, A., et al., Genes to Cells, 1, 569-579 (1996).
28. Sidhu, M.S., et al., Virology, 208, 800-807 (1995).
29. Shaffer, M.F., et al., J. Immunol., 41, 241-256 (1941).
30. Enders, J.F., et al., N. Engl. J. Med., 263, 153-159 (1960).
31. Enders, J.F., and Peebles, M.E., Proc. Soc. Exp. Biol. Med., 86, 227-286 (1954).
32. Schwarz, A.J.F., Am. J. Dis. Child., 103, 216-219 (1962).
33. Griffin, D.E., and Bellini, W.J., pages 1267-1312 of Volume 1, Fields Virology, B.N. Fields, et al., Eds. (3rd ed., Raven Press, 1996).

SUBSTITUTE SHEET (RULE 26)

- 82 -

34. Birrer, M.J., et al., Virology, 108, 381-390 (1981).
35. Birrer, M.J., et al., Nature, 293, 67-69 (1981).
36. Norby, E., et al., pages 481-507, in The Paramyxoviruses, D. Kingsbury, Ed. (Plenum Press, 1991).
37. Peebles, M.E., pages 427-456, in The Paramyxoviruses, D. Kingsbury, Ed. (Plenum Press, 1991).
38. Egelman, E.H., et al., J. Virol., 63, 2233-2243 (1989).
39. Udem, S.A., et al., J. Virol. Methods, 8, 123-136 (1984).
40. Udem, S.A., and Cook, K.A., J. Virol., 49, 57-65 (1984).
41. Moyer, S.A., and Horikami, S.M., pages 249-274, in The Paramyxoviruses, D. Kingsbury, Ed. (Plenum Press, 1991).
42. Blumberg, B., et al., pages 235-247, in The Paramyxoviruses, D. Kingsbury, Ed. (Plenum Press, 1991).
43. Berrett, T., et al., pages 83-102, in The Paramyxoviruses, D. Kingsbury, Ed. (Plenum Press, 1991).
44. Tordo, N., et al., Sem. in Virology, 3, 341-357 (1992).
45. Cattaneo, R., et al., EMBO J., 6, 681-688 (1987).
46. Crowley, J.C., et al., Virology, 164, 498-506 (1988).
47. Banerjee, A.K., and Barik, S., et al., Virology, 188, 417-428 (1992).
48. Castaneda, S.J., and Wong, T.C., J. Virol., 63, 2977-2986 (1989).

SUBSTITUTE SHEET (RULE 26)

49. Chan, J., et al., pages 221-231, in Genetics and Pathogenicity of Negative Stranded Viruses, B.W.J. Mahy and D. Kolakofsky, Eds. (Elsevier Biomedical Press, 1989).
50. Blumberg, B., et al., Cell, 23, 837-845 (1981).
51. Blumberg, B., et al., Cell, 32, 559-567 (1983).
52. Kolakofsky, D., and Blumberg, B.M., pages 203-213, in Virus Persistence, B.M.J. Mahy, et al., Eds. (Cambridge University Press, 1982).
53. Castaneda, S.J., and Wong, T.C., J. Virol., 64, 222-230 (1990).
54. Curran, J.A., and Kolakofsky, D., Virology, 182, 168-176 (1991).
55. Sidhu, M.S., et al., Virology, 193, 66-72 (1993).
56. Sidhu, M.S., et al., Virology, 202, 631-641 (1994).
57. Collins, P.L., et al., pages 1205-1241 of Volume 1, Fields Virology, B.N. Fields, et al., Eds. (3rd ed., Raven Press, 1996).
58. Crookshanks, F.K., and Belshe, R.B., J. Med. Virol., 13, 243-249 (1984).
59. Crookshanks-Newman, F.K., and Belshe, R.B., J. Med. Virol., 18, 131-137 (1986).
60. Hall, S.L., et al., Virus Res., 22, 173-184 (1992).
61. Karron, R.A., et al., J. Inf. Dis., 172, 1445-1450 (1995).
62. Anderson, L.J., et al., J. Infect. Dis., 151, 626-633 (1985).
63. Collins, P.L., pages 103-162 of The Paramyxoviruses, D.W. Kingsbury, Ed. (Plenum Press, NY and London, 1991).

- 84 -

64. Sullender, W.M., J. Virology, 65, 5425-5434 (1991).
65. Lerch, R.A., et al., J. Virology, 64, 5559-5569 (1990).
66. Mallipeddi, S.K., and Samal, S.K., J. Gen Virol., 74, 2787-2791 (1993).
67. Johnson, P.R., et al., J. Virology, 61, 3163-3166 (1987).
68. Stott, E.J., et al., J. Virology, 61, 3855-3861 (1987).
69. Henderson, F.W., et al., N. Engl. J. Med., 300, 530-534 (1979).
70. Hall, S.L., et al., J. Infect. Dis., 163, 693-698 (1991).
71. Mufson, M.A., et al., J. Gen. Virol., 66, 2111-2124 (1985).
72. Glezen, W.P., et al., Am. J. Dis. Child., 140, 543-546 (1986).
73. Hemming, V.G., et al., Clin. Microbiol. Res., 8, 22-33 (1995).
74. Collins, P. L. et. al., pages 1313-1351 of volume 1, Fields Virology, B. N. Fields, et al., Eds. (3rd ed., Raven Press, 1996).
75. Ling, R., and Pringle, C.R., J. Gen. Virol., 70, 1427-1440 (1989).
76. Yu, Q., et al., J. Virology, 69, 2412-2419 (1995).
77. McIntosh, K., and Chanock, R.M., pages 1045-1072 of Virology, B.N. Fields, et al., Eds. (2nd ed., Raven Press, 1990).
78. Heminway, B.R., et al., page 167 of Abstracts of the IX International Congress of Virology, P17-2, (1993).
79. Mink, M.A., et al., Virology, 185, 615-624 (1991).

SUBSTITUTE SHEET (RULE 26)

- 85 -

80. Dickens, L.E., et al., J. Virology, 52, 364-369 (1990).

81. Wagner, R.R., and Rose, J.K., pages 1121-1135 of volume 1, Fields Virology, B.N. Fields, et al., Eds. (3rd ed., Raven Press, 1996).

82. Barik, S., J. Gen. Virol., 74, 485-490 (1993).

83. Collins, P.L., et al., pages 259-264 of Vaccines 93: modern approaches to new vaccines including prevention of AIDS, F. Brown et al., Eds. (Cold Spring Harbor Laboratory Press, NY, 1993).

84. Kuo, L., et al., J. Virology, 70, 6892-6901 (1996).

85. Huang, Y.T., and Wertz, G.W., J. Virology, 43, 150-157 (1982).

86. Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

87. Ray, R., et al., J. Virol., 69, 1959-1963 (1995).

88. Ray, R., et al., J. Virol., 70, 580-584 (1996).

89. Stokes, A., et al., Virus Research, 30, 43-52 (1993).

90. U.S. Patent Application No. 08/059,444.

- 86 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Udem, Stephen A.  
Sidhu, Mohinderjit S.  
Tatem, Joanne M.  
Murphy, Brian R.  
Randolph, Valerie B.

(ii) TITLE OF INVENTION: 3' Genomic Promoter Region and  
Polymerase Gene Mutations Responsible for Attenuation in  
Viruses of the Order Designated Mononegavirales

(iii) NUMBER OF SEQUENCES: 79

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: American Home Products Corporation  
(B) STREET: One Campus Drive  
(C) CITY: Parsippany  
(D) STATE: New Jersey  
(E) COUNTRY: United States  
(F) ZIP: 07054

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US  
(B) FILING DATE:  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Gordon, Alan M.  
(B) REGISTRATION NUMBER: 30,637  
(C) REFERENCE/DOCKET NUMBER: 33,294 PCT

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 973/683-2157  
(B) TELEFAX: 973/683-4117

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15894 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single

- 87 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACCAAACAAA GTTGGGTAAG GATAGATCAA TCAATGATCA TATTCTAGTG CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCTCTCA	240
TTACCACTCG ATCCAGACTT CTGGACCGGT TGGTCAGGTT AATTGGAAC CCGGATGTGA	300
GCGGGCCCCA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAATTAG TAGTGATCAA TCCAGGTTCTG	540
GATGGTTCTGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCCCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGAACACCCG	840
GAAACAAACC CAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GA CTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAC CTTTACCAGC	1020
AAATGGGGGA AACTGCACCC TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAGTTCA	1080
GTGCAGGATC ATACCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGCCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTAAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCTGAACTCG	1260
GTATCACTGC CGAGGATGCA AGGCTTGTTT CAGAGATTGC AATGCATACT ACTGAGGACA	1320

SUBSTITUTE SHEET (RULE 26)

- 88 -

AGATCAGTAG	AGCGGTTGGA	CCCAGACAAG	CCCAAGTATC	ATTTCTACAC	GGTGATCAAA	1380
GTGAGAATGA	GCTACCGAGA	TTGGGGGGCA	AGGAAGATAG	GAGGGTCAAA	CAGAGTCGAG	1440
GAGAAGCCAG	GGAGAGCTAC	AGAGAAACCG	GGCCCAGCAG	AGCAAGTGAT	GCGAGAGCTG	1500
CCCATCTTCC	AACCGGCACA	CCCCTAGACA	TTGACACTGC	ATCGGAGTCC	AGCCAAGATC	1560
CGCAGGACAG	TCGAAGGTCA	GCTGACGCCC	TGCTTAGGCT	GCAAGCCATG	GCAGGAATCT	1620
CGGAAGAACA	AGGCTCAGAC	ACGGACACCC	CTATAGTGTA	CAATGACAGA	AATCTTCTAG	1680
ACTAGGTGCG	AGAGGCCGAG	GACCAGAACA	ACATCCGCCT	ACCCTCCATC	ATTGTTATAA	1740
AAAACCTTAGG	AACCAGGTCC	ACACAGCCGC	CAGCCCATCA	ACCATCCACT	CCCACGATTG	1800
GAGCCGATGG	CAGAAGAGCA	GGCACGCCAT	GTCAAAAACG	GACTGGAATG	CATCCGGGCT	1860
CTCAAGGCCG	AGCCCATCGG	CTCACTGGCC	ATCGAGGAAG	CTATGGCAGC	ATGGTCAGAA	1920
ATATCAGACA	ACCCAGGACA	GGAGCGAGCC	ACCTGCAGGG	AAGAGAAGGC	AGGCAGTTCG	1980
GGTCTCAGCA	AACCATGCCT	CTCAGCAATT	GGATCAACTG	AAGGCGGTGC	ACCTCGCATC	2040
CGCGGCCAGG	GACCTGGAGA	GAGCGATGAC	GACGCTGAAA	CTTTGGGAAT	CCCCCAAGA	2100
AATCTCCAGG	CATCAAGCAC	TGGGTTACAG	TGTTATTATG	TTTATGATCA	CAGCGGTGAA	2160
GCGGTTAAGG	GAATCCAAGA	TGCTGACTCT	ATCATGGTTC	AATCAGGCCT	TGATGGTGAT	2220
AGCACCTCT	CAGGAGGAGA	CAATGAATCT	GAAAACAGCG	ATGTGGATAT	TGGCGAACCT	2280
GATACCGAGG	GATATGCTAT	CACTGACCGG	GGATCTGCTC	CCATCTCTAT	GGGGTTCAGG	2340
GCTTCTGATG	TTGAAACTGC	AGAAGGAGGG	GAGATCCACG	AGCTCCTGAG	ACTCCAATCC	2400
AGAGGCAACA	ACTTTCCGAA	GCTTGGGAAA	ACTCTCAATG	TTCCTCCGCC	CCCGGACCCC	2460
GGTAGGGCCA	GCACTTCCGA	GACACCCATT	AAAAAGGGCA	CAGACGCGAG	ATTAGCCTCA	2520
TTTGGAACGG	AGATCGCGTC	TTTATTGACA	GGTGGTGCAA	CCCAATGTGC	TCGAAAGTCA	2580
CCCTCGGAAC	CATCAGGGCC	AGGTGCACCT	GCGGGGAATG	TCCCCGAGTG	TGTGAGCAAT	2640
GCCGCACTGA	TACAGGAGTG	GACACCCGAA	TCTGGTACCA	CAATCTCCCC	GAGATCCCAG	2700
AATAATGAAG	AAGGGGGAGA	CTATTATGAT	GATGAGCTGT	TCTCTGATGT	CCAAGATATT	2760
AAAACAGCCT	TGGCCAAAAT	ACACGAGGAT	AATCAGAAGA	TAATCTCCAA	GCTAGAATCA	2820
CTGCTGTTAT	TGAAGGGAGA	AGTTGAGTCA	ATTAAGAAGC	AGATCAACAG	GCAAAATATC	2880

SUBSTITUTE SHEET (RULE 26)



- 89 -

AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG	2940
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA	3000
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCGTTGC CAGCCGACAA	3060
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTACAG	3120
CTAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCTTGA CACCGGCCCT	3180
GCATCACGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG	3240
CGTTACCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG	3300
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG	3360
CCAGTCGACC CAACTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT	3420
GCCTCCCAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA	3480
AAGGGTCGAT CGCTCCGATA CAACCCACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG	3540
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTATG TACATGTTTC	3600
TGCTGGGGGT TGTTGAGGGC AGCGATCCCC TAGGGCCTCC AATCGGGCGA GCATTGGGT	3660
CCCTGCCCTT AGGTGTTGGC AGATCCACAG CAAAGCCCGA AGAACTCCTC AAAGAGGCCA	3720
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGA AAAACTG GTGTTCTACA	3780
ACAACACCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT	3840
TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTCGATACC CCGCAGAGGT	3900
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTCTTA	3960
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCCTTA	4020
GGATTGACAA GGCATAGGC CCTGGGAAGA TCATCGACAA TACAGAGCAA CTTCTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCACCTGGT GGGATAGGGG	4200
GCACCACTCT TCACATTAGA AGCACAGGCA AGATGAGCAA GACTCTCCAT GCACAACTCG	4260
GGTTCAAGAA GACCTTATGT TACCCGCTGA TGGATATCAA TGAAGACCTT AATCGATTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTGTCAGCCA TCAGTTCTCT	4380
AAGAATTCCG CATTTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC	4440

SUBSTITUTE SHEET (RULE 26)

- 90 -

TGTAGACCGT AGTGCCGAGC AATGCCCGAA AACGACCCCC CTCACAATGA CAGCCAGAAG	4500
GGCCGGACAA AAAAGCCCCC TCCGAAAGAC TCCACGGACC AAGCGAGAGG CCAGCCAGCA	4560
GGCGACGGCA AGCGCGAACA CCAGGCGGCC CCAGCACAGA ACAGCCCTGA CACAAGGCCA	4620
CCACCAGCCA CCCCAATCTG CATCCTCCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGC	4680
TGCCCCGAT CCAAACCACC AACC GCATCC CCACCACCCC CGGGAAAGAA ACCCCCAGCA	4740
ATTGGAAGGC CCCTCCCCCT CTTCTCAAC ACAAGAACTC CACAACCGAA CCGCACAAAGC	4800
GACCGAGGTG ACCCAACCGC AGGCATCCGA CTCCCTAGAC AGATCCTCTC TCCCCGGCAA	4860
ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAC AGAACCAGA CCCC GGCCCA	4920
CGGCGCCGCG CCCCCAACCC CCGACAACCA GAGGGAGCCC CCAACCAATC CCGCCGGTTC	4980
CCCCGGTGCC CACAGGCAGG GACACCAACC CCCGAACAGA CCCAGCACCC AACCATCGAC	5040
AATCCAAGAC GGGGGGGCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCACCGC	5100
GAGGAAGCCC ACCCACCCTA CACACGACCA CGGCAACCAA ACCAGAACCC AGACCACCT	5160
GGGCCACCAG CTCCAGACT CGGCCATCAC CCCGCAGAAA GGAAAGGCCA CAACCCGCGC	5220
ACCCAGCCC CGATCCGGCG GGGAGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC	5280
CGAAGGACCC CCGAACCGCA AAGGACATCA GSTATCCACA GCCTCTCCAA GTCCCCGGT	5340
CTCTCTTT TCTCGAAGGG ACCAAAAGAT CAATCCACCA CACCCGACGA CACTCAACTC	5400
CCCACCCCTA AAGGAGACAC CGGGAATCCC AGAATCAAGA CTCATCCAAT GTCCATCATG	5460
GGTCTCAAGG TGAACGTCTC TGCCATATTC ATGGCAGTAC TGTAACTCT CCAGACACCC	5520
ACCGGTCAA TCCATTGGGG CAATCTCTCT AAGATAGGGG TGGTAGGAAT AGGAAGTGCA	5580
AGCTACAAAG TTATGACTCG TTCCAGCCAT CAATCATTAG TCATAAAATT AATGCCCAAT	5640
ATAACTCTCC TCAATAACTG CACGAGGGTA GAGATTGCAG AATACAGGAG ACTACTGAGA	5700
ACAGTTTTGG AACCAATTAG AGATGCACTT AATGCAATGA CCCAGAATAT AAGACCGGTT	5760
CAGAGTGTAG CTCAAGTAG GAGACACAAG AGATTTGCGG GAGTAGTCCT GGCAGGTGCG	5820
GCCCTAGGCG TTGCCACAGC TGCTCAGATA ACAGCCGGCA TTGCACTTCA CCAGTCCATG	5880
CTGAACTCTC AAGCCATCGA CAATCTGAGA GCGAGCCTGG AAAC TACTAA TCAGGCAATT	5940
GAGGCAATCA GACAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC	6000

SUBSTITUTE SHEET (RULE 26)

- 91 -

ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTTAAT CGGCCAGAAG	6060
CTCGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CATTATTTGG CCCCAGCTTA	6120
CGGGACCCCA TATCTGCGGA GATATCTATC CAGGCTTTGA GCTATGCGCT TGGAGGAGAC	6180
ATCAATAAGG TGTTAGAAAA GCTCGGATAC AGTGGAGGTG ATTTACTGGG CATCTTAGAG	6240
AGCAGAGGAA TAAAGGCCCG GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTCCTC	6300
AGTATAGCCT ATCCGACGCT GTCCGAGATT AAGGGGGTGA TTGTCCACCG GCTAGAGGGG	6360
GTCTCGTACA ACATAGGCTC TCAAGAGTGG TATACCACTG TGCCCAAGTA TGTTGCAACC	6420
CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT	6480
GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TCCACCAAGT CCTGTGCTCG TACACTCGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCC	6720
GTAGTCGAGG TGAACGGCGT GACCATCCAA GTCGGGAGCA GGAGGTATCC AGACGCTGTG	6780
TACTTGCA CA GAATTGACCT CGGTCTCTCC ATATCATTGG AGAGGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTAGCA TAGTCTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACGGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCACCCA GCATCAAGCC CACCTGAAAT	7200
TATCTCCGGC TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATTAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA	7320
TCCCAAGGGA AGTAGGATAG TCATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTGTTT TCATGTCTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500
TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560

SUBSTITUTE SHEET (RULE 26)

- 92 -

AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTCACTGACC TAGTGAAATT 7620  
CATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC 7680  
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT 7740  
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC 7800  
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA 7860  
ATTCTCAAAC ATGTCGCTGT CCCTGTTAGA CTTGTATTTA AGTCGAGGTT ACAATGTGTC 7920  
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTATGGGGGA ACTTACCTAG TGGAAAAGCC 7980  
TAATCTGAGC AGCAAAAGGT CAGAGTTGTC ACAACTGAGC ATGTACCGAG TGTTTGAAGT 8040  
AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATCTTGA 8100  
GCAACCAGTC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAACCTGC 8160  
AGCCCTTTGT CACGGGGAAG ATTCTATCAC AATTCCTTAT CAGGGATCAG GGAAAGGTGT 8220  
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT 8280  
CCCCTTATCA ACGGATGATC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT 8340  
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGAACAGATG ACAAGTTGCG 8400  
AATGGAGACA TGCTTCCAAC AGGCGTGTA GGGTAAATC CAAGCACTCT GCGAGAATCC 8460  
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTGATCT 8520  
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA 8580  
CGGTTCAGGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGAATATCCC 8640  
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA 8700  
GGTTAGTCCC AACCTCTTCA CTGTCCCAAT TAAGGAAGCA GGCGAAGACT GCCATGCCCC 8760  
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAATC AGTTCCAATC TGGTGATTCT 8820  
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCAGGG TTGAACATGC 8880  
TGTGGTTTAT TACGTTTACA GCCCAGGCCG CTCATTTTCT TACTTTTATC CTTTATAGTT 8940  
GCCTATAAAG GGGGTCCCCA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT 9000  
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC 9060  
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATCGCAG 9120

SUBSTITUTE SHEET (RULE 26)

- 93 -

ATAGGGCTGC TAGTGAACCA ATCACATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACC GCC TAAAAAACGG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCCACT	9480
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540
CGAGGAAGAT CCGTGAATC CTCAAAAAGG GGAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCTTAAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAAC TTG GAGTTTACAT GCACAGCTCC CAGTGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACATTTGA ACTGGTTTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTCACTTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCATGAG TTAATTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA	10260
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT	10560
TTGCTGGAGT GAAATTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT	10620
ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG	10680

SUBSTITUTE SHEET (RULE 26)

AGTTCCTGCG TTACGACCCT CCCAAGGGAA CCGGGTCACG GAGGCTTGTA GATGTTTTCC 10740  
TTAATGATTC GAGCTTTGAC CCATATGATG TGATAATGTA TGTGTAAAGT GGAGCTTACC 10800  
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG 10860  
GTAGACTTTT TGCTAAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC 10920  
TAATCTCAAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCACG 10980  
ATTTGACTAA GGCACTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA 11040  
GTCACAGGGG GGGGCCAGTC TTA AAAACCT ACTCCGAAG CCCAGTCCAC ACAAGTACCA 11100  
GGAACGTGAG AGCAGCAAAA GGGTTTATAG GGTTCCTCA AGTAATTCGG CAGGACCAAG 11160  
ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACAGACAGT CAGTGCATTT ATCAGGACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTG CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCTGTA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCAT ATCCCGTTAT 11400  
ATAAAGTCCC CAATGATCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATTG CCTATCTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCCTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCTCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CTTGTCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTCTGATCT 11940  
CTCTTGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCCTCACTAA TGCCTGAAGA GACCCTCCAT CAAGTAATGA 12180  
CACAAACAACC GGGGGACTCT TCATTCCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240

- 95 -

TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA	12300
TCCATAGTCC AAACCCAATG TTAAAAGGAT TATTCCATGA TGACAGTAAA GAAGAGGACG	12360
AGGGACTGGC GGCATTCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA	12420
TCCTGGATCA TAGTGTCA CA GGGGCAAGAG AGTCTATTGC AGGCATGCTG GATACCACAA	12480
AAGGCCTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT	12540
TGTCCAATTA TGA CTATGAA CAATTCAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA	12600
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GGCGAGAGCT CTAAGAAGCC	12660
ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC	12720
TAGAATCTAT GCGAGGCCAC CTTATTCGGC GTCATGAGAC ATGTGTCATC TGCGAGTGTG	12780
GATCAGTCAA CTACGGATGG TTTTGTGCC CCTCGGGTTG CCAACTGGAT GATATTGACA	12840
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA	12900
TGAAGCTTGC CTTTCGTAAGA GCCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA	12960
CAGTGACTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG	13020
CTAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT	13080
CGACTAATTT AGCGCATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT	13140
CCCTTGTCGG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT	13200
CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAAGG AATGCTTCTA GGGTTGGGTG	13260
TTT TAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC	13320
TTCACGTCGA AACAGATTGT TCGTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT	13380
CCC GCAAGCT AGAGCTGAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC	13440
CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG	13500
AATTGTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT	13560
CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA	13620
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTCATAGAG CCAAGATTAT	13680
TCACTATCTA CTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA	13740
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTTGTCATC GTTCCTTTCT AGAATGAGCA	13800

SUBSTITUTE SHEET (RULE 26)

- 96 -

AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT	13860
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGCACA	13920
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG	13980
AAGAGTTAGA AGAGTTCACA TTTCTCTTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA	14040
GATTCGACAA CATCCAGGCA AAACACTTAT GTGTTCTGGC AGATTGTGAC TGTCAACCAG	14100
GGACCTGCCC ACCAATTCTGA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC	14160
ATATCAAGGC AGAGGCTAGG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA	14220
TTGTAGACCA TTAATCATGC TCTCTGACTT ATCTCCGGCG AGGATCGATC AAACAGATAA	14280
GATTGAGAGT TGATCCAGGA TTCATTTTCG ACGCCCTCGC TGAGGTAAAT GTCAGTCAGC	14340
CAAAGATCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGATTTTCTG CCCCCACACG	14400
ATGATGTTGC AAAATTGCTC AAAGATATCA ACACAAGCAA GCACAATCTT CCCATTTTCAG	14460
GGGGCAATCT CGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTG AACTCATCTG	14520
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAAG	14580
ACGGCTTGTT CTTGGGTGAG GGATCGGGTT CTATGTTGAT CACTTATAAG GAGATACTTA	14640
AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG	14700
AATTAGCACC CTATCCCTCC GAAGTTGGCC TTGTGCAACA CAGAATGGGA GTAGGTAATA	14760
TTGTCAAAGT GCTCTTTAAC GGGAGGCCCG AAGTCACGTG GGTAGGCAGT GTAGATTGCT	14820
TCAATTTTCAT AGTTAGTAAT ATCCCTACCT CTAGTGTGGG GTTTATCCAT TCAGATATAG	14880
AGACCTTGCC TAACAAAGAT ACTATAGAGA AGCTAGAGGA ATTGGCAGCC ATCTTATCGA	14940
TGGCTCTGCT CCTGGGCAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTCAGCG	15000
GGGATTTTGT TCAGGGATTT ATAAGTTATG TAGGGTCTCA TTATAGAGAA GTGAACCTTG	15060
TATACCCTAG ATACAGCAAC TTCATATCTA CTGAATCTTA TTTGGTTATG ACAGATCTCA	15120
AGGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGA	15180
GGACTTCACC TGGACTTATA GGTCACATCC TATCCATTAA GCAACTAAGC TGCATACAAG	15240
CAATTGTGGG AGACGCAGTT AGTAGAGGTG ATATCAATCC TACTCTGAAA AAACCTACAC	15300
CTATAGAGCA GGTGCTGATC AATTGCGGGT TGGCAATTAA CGGACCTAAG CTGTGCAAA	15360

SUBSTITUTE SHEET (RULE 26)



- 97 -

AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAATTCT ATACTCATCC 15420  
 TCTACAGGGA GTTGGCAAGA TTCAAAGACA ACCAAAGAAG TCAACAAGGG ATGTTCCACG 15480  
 CTTACCCCGT ATTGGTAAGT AGCAGGCAAC GAGAACTTAT ATCTAGGATC ACCCGCAAAT 15540  
 TTTGGGGGCA CATTCTTCTT TACTCCGGGA ACAGAAAGTT GATAAATAAG TTTATCCAGA 15600  
 ATCTCAAGTC CGGCTATCTG ATACTAGACT TACACCAGAA TATCTTCGTT AAGAATCTAT 15660  
 CCAAGTCAGA GAAACAGATT ATTATGACGG GGGGTTTGAA ACGTGAGTGG GTTTTTAAGG 15720  
 TAACAGTCAA GGAGACCAAA GAATGGTATA AGTTAGTCGG ATACAGTGCC CTGATTAAGG 15780  
 ACTAATTGGT TGAAGTCCGG AACCTAATC CTGCCCTAGG TGGTTAGGCA TTATTTGCAA 15840  
 TATATTAAAG AAAACTTTGA AAATACGAAG TTCTATTCC CAGCTTTGTC TGGT 15894

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
 1 5 10 15  
 Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
 20 25 30  
 Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
 35 40 45  
 Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
 50 55 60  
 Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
 65 70 75 80  
 Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn  
 85 90 95

SUBSTITUTE SHEET (RULE 26)

- 98 -

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys  
 100 105 110  
 Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu  
 115 120 125  
 Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp  
 130 135 140  
 Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln  
 145 150 155 160  
 Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg  
 165 170 175  
 Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr  
 180 185 190  
 Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
 195 200 205  
 Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
 210 215 220  
 Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
 225 230 235 240  
 Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly  
 245 250 255  
 Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
 260 265 270  
 Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu  
 275 280 285  
 Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe  
 290 295 300  
 Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly  
 305 310 315 320  
 Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr  
 325 330 335  
 Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe  
 340 345 350  
 Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu  
 355 360 365  
 Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr

SUBSTITUTE SHEET (RULE 26)

- 99 -

370	375	380
Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr		
385	390	395 400
Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His		
	405	410 415
Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr		
	420	425 430
His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe		
	435	440 445
Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu		
	450	455 460
Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr		
	465	470 475 480
Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg		
	485	490 495
Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp		
	500	505 510
Val Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe		
	515	520 525
Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg		
	530	535 540
Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala		
	545	550 555 560
Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly		
	565	570 575
Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala		
	580	585 590
Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro		
	595	600 605
Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn		
	610	615 620
Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Val Ile Arg Gln		
	625	630 635 640
Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val		
	645	650 655

SUBSTITUTE SHEET (RULE 26)

- 100 -

Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670  
 Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly  
 675 680 685  
 Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val  
 690 695 700  
 Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile  
 705 710 715 720  
 Pro Leu Tyr Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met  
 725 730 735  
 Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile  
 740 745 750  
 Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser  
 755 760 765  
 Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro  
 770 775 780  
 Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr  
 785 790 795 800  
 Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His  
 805 810 815  
 His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr  
 820 825 830  
 Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845  
 Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860  
 Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880  
 Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895  
 Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910  
 Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile  
 915 920 925

SUBSTITUTE SHEET (RULE 26)

- 101 -

Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940

Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960

Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu  
 965 970 975

Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990

Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005

Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His  
 1010 1015 1020

Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
 1025 1030 1035 1040

Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
 1045 1050 1055

Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
 1060 1065 1070

Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
 1075 1080 1085

Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
 1090 1095 1100

Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly  
 1105 1110 1115 1120

Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu  
 1125 1130 1135

Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg  
 1140 1145 1150

Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly  
 1155 1160 1165

His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser  
 1170 1175 1180

Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp  
 1185 1190 1195 1200

Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr

SUBSTITUTE SHEET (RULE 26)

- 102 -

1205	1210	1215
Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser 1220	1225	1230
Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala 1235	1240	1245
Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg 1250	1255	1260
Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile 1265	1270	1275 1280
Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln 1285	1290	1295
Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr 1300	1305	1310
Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp 1315	1320	1325
Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu 1330	1335	1340
Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val 1345	1350	1355 1360
Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp 1365	1370	1375
His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu 1380	1385	1390
Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp 1395	1400	1405
Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe 1410	1415	1420
Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr 1425	1430	1435 1440
Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met 1445	1450	1455
Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile 1460	1465	1470
Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly 1475	1480	1485

SUBSTITUTE SHEET (RULE 26)

- 103 -

Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro  
1490 1495 1500

Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg  
1505 1510 1515 1520

Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro  
1525 1530 1535

Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His  
1540 1545 1550

Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met  
1555 1560 1565

Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu  
1570 1575 1580

Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val  
1585 1590 1595 1600

Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala  
1605 1610 1615

Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg  
1620 1625 1630

Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala  
1635 1640 1645

Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val  
1650 1655 1660

Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
1665 1670 1675 1680

Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
1685 1690 1695

Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
1700 1705 1710

Met Ser Ile Lys Asp Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
1715 1720 1725

Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
1730 1735 1740

Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn  
1745 1750 1755 1760

SUBSTITUTE SHEET (RULE 26)

- 104 -

Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
 1765 1770 1775  
 Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
 1780 1785 1790  
 Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
 1795 1800 1805  
 Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
 1810 1815 1820  
 Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
 1825 1830 1835 1840  
 Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp  
 1845 1850 1855  
 Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr  
 1860 1865 1870  
 Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys  
 1875 1880 1885  
 Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
 1890 1895 1900  
 Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
 1905 1910 1915 1920  
 Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His  
 1925 1930 1935  
 Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser  
 1940 1945 1950  
 Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met  
 1955 1960 1965  
 Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr  
 1970 1975 1980  
 Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys  
 1985 1990 1995 2000  
 Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro  
 2005 2010 2015  
 Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly  
 2020 2025 2030  
 Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp

SUBSTITUTE SHEET (RULE 26)



- 105 -

2035	2040	2045
Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr 2050	2055	2060
Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met 2065	2070	2075 2080
Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile 2085	2090	2095
Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly 2100	2105	2110
Asn Arg Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr 2115	2120	2125
Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys 2130	2135	2140
Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val 2145	2150	2155 2160
Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly 2165	2170	2175
Tyr Ser Ala Leu Ile Lys Asp 2180		

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ACCAAACAAA GTTGGGTAAG GATAGATCAA TCAATGATCA TATTCTAGTA CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTC	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCGGGA GATTCTCTCA	240

SUBSTITUTE SHEET (RULE 26)

- 106 -

TTACCACTCG ATCTAGACTT CTGGACCGGT TGGTCAGGTT AATTGGAAC CCGGATGTGA	300
GCGGGCCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATATTTT TCACATGATG ATCCAAGTAG TAGTGATCAA TCCAGGTTCTG	540
GATGGTTCTGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCTCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGAACACCCG	840
GGAACAAACC CAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GACTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAT CTTTACCAGC	1020
AAATGGGGGA AACTGCACCA TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAGTTCA	1080
GTGCAGGATC ATACCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGCCGAT CTTACTTCGA TCCAGCATAT TTCAGACTAG	1200
GGCAAGAGAT GGTGAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCTGAACTCG	1260
GTATCACTGC CGAAGATGCA AGGCTTGTTT CAGAGATCGC AATGCATACT ACAGAGGACA	1320
GGATCAGTAG AGCGGTTGGA CCCAGACAAT CCCAAGTGTC ATTCCTACAC GGTGATCAAA	1380
ATGAAAATGA GCTACCGAGA TGGGGGGGTA AGGAAGATAT GAGGGTCAA CAGAGTCGGG	1440
GAGAAGCCAG AGAGAGCTAC AGAGAAACCA GGCCAGCAG AGCAAGTGAC GCGAGAGCTA	1500
CCCATCCTCC AACCGACACA CCCTTAGACA TTGACACTGC ATCGGAGTCC AGCCAAGATC	1560
CGCAGGACAG TCGAAGGTCA GCTGACGCCC TGCTCAGGCT GCAAGCCATG GCAGGAATCT	1620
CGGAAGAACA AGGCTCAGAC ACGGACACCC CTAGAGTGTA CAATGACAGA GATCTTCTAG	1680
ACTAGGTGCA AGAGGCCGAG GACCAGAACA ACATCCGCCT ACCCTCCATC ATTGTTATAA	1740
AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCCACCA ACCATCCACT CCCACGATTG	1800

SUBSTITUTE SHEET (RULE 26)

- 107 -

GGGCCGATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGGCT	1860
CTCAAGGCCG AGCCCATCGG CTCACTGGCC ATCGAGGAAG CTATGGCAGC ATGGTCAGAA	1920
ATATCAGACA ACCCAGGACA GGAGCGAGCC GCCTGCAAGG AAGAGAAGGC AAGCAGTCCG	1980
GGTCTCAGCA AACCATGCCT CTCAGCAATT GGATCAACTG AAGGCGGTGC ACCTCGCATC	2040
CGCGGTCAGG GATCTGGAGA GAGCGATGAC GACGCTGAAA CTTTGGGAAT CCCCTCAGGA	2100
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTATTATG TTTATGATCA CAGCGGTGAA	2160
GCGGTAAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT	2220
AGCACCTCT CAGGAGGAGA CAATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT	2280
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTTCAGG	2340
GCTTCTGATG TTGAAACTGC AGAAGGAGGG GAGATCCACG AGCTCCTGAG ACTCCAATCC	2400
AGAGGCAACA ACTTTCCAAA GCTTAGGAAA ACTCTCAATG TTCCCCCGCC CCCGGACCCT	2460
GGTAGGGCCA GCACTTCCGA GACACCCATT AAAAAGGGCA CAGACGCGAG ATTAGCCTCA	2520
TTTGAACGG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA	2580
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GCGGGGAATG TCCCCGAGTG TGTGAGCAAT	2640
GCCGTACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCCAG	2700
AATAATGAAG AAGGGGGAGA TTATTATGAT GATGAGCTGT TCTCTGATGT CCAAGATATT	2760
AAAAACGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCACCAA GCTAGAATCA	2820
CTGCTGTTAT TGAAGGGGGA AGTTGAGTCA ATCAAGAAGC AGATCAACAG GCAAAATATC	2880
AGCATATCCA CCTTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG	2940
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA	3000
GGCAGAGATT CAGGCCGAGC ACTGGCTGAA GTTCTCAAGA AACCCTTGC CAGCCGACAA	3060
ATCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTAG	3120
CTAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCCGGA CACCGGCCCT	3180
GCATCAGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG	3240
CGTTACCTGA TGAATCTCCT TGATGACATC AAAGGAGCCA ACGATCTTGC CAAGTTCCAC	3300
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG	3360

SUBSTITUTE SHEET (RULE 26)

- 108 -

CCAGTCGACC TAGCTAATAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT 3420  
GCCTCCCAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA 3480  
AAGGGTCGAT CGCTCCGATA CAACCCACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG 3540  
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GAAAAGATGA ATGTTTTATG TACATGTTTC 3600  
TGCTGGGGGT TGTGAGGAC AGCGATCTCC TAGGGCCTCC AATCGGGCGA GCATTTGGGT 3660  
CTCTGCCCTT AGGTGTTGGC AGATCCACAG CAAAACCCGA AGAACTCCTC AAAGAGGCCA 3720  
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAAACTG GTGTTCTACA 3780  
ACAACACCCC ACTAACTCTC CTCATACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT 3840  
TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTGGATACC CCGCAGAGGT 3900  
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCAGATAA CGGGTATTAC ACCGTTTCCTA 3960  
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCTTA 4020  
GGATTGACAA GGCGATTGGC CATGGGAAGA TCATCGACAA TGCAGAGCAA CTTCTGAGG 4080  
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA AAGTGAAGTC TACTCTGCCG 4140  
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCACCTGGT GGGATAGGGG 4200  
GCACCAGTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACTCG 4260  
GGTTCAAGAA GACCCTATGT TACCCACTGA TGGATATCAA TGAAGACCTT AATCGATTAC 4320  
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC 4380  
AAGAATTCCG CATTTACGAC GACGTTATCA TAAATGATGA CCAAGGATTA TTCAAAGTTC 4440  
TGTAGACCGT AGTGCCAGC AATGCCCCGAA GACGACCCTC CTCACAATGA CAGCCAGAAG 4500  
GCCCCGAAAA AAAGGCCCCC TCCGAAAGAC TCCACAGACC AAATGAGAGG CCAGCCAGCA 4560  
GCTGACGGCA AGCACGAACA CCAGGCGGCC CCAGCACAGA ACAGCCCTGA CATAAGGCCA 4620  
CCACCAGCCA TCCCAATCTG CATCCTCCTC GTAGGACCCC CGAGGACCAA CCCCCAAGGT 4680  
TGCCCCCCAC CCAAACCACC AACCGCATCC CTACCACCCC CGGGAAAGAA ACCCCCAGCA 4740  
ACTGGAAGAG CCCTTCCCCT TTCCCTCAAC ACAAGAACTC CACAACCGAA CCACACAAGC 4800  
GACCGAGGTG ACCCAACCGC AGGCACCCGA CTCCCTAGAC AGATCCTCTC CCCCTGGCAA 4860  
ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAC AGAACCAGCA CCCCCGCCCA 4920

SUBSTITUTE SHEET (RULE 26)

- 109 -

CGGCGCCGCG CCCCCAACCC CCGACAACCA GAGGGAGCCC CCAACCAATC CCGCCGGCTC 4980  
CCCCGGTGCC CACAGGCAGG CACACCAACC CCCGAACAGA CCCAGCACCC AGCCATCGAC 5040  
AATCCAAGAC GGGGGGGCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCACCGC 5100  
GAGGAAGCCC ACCCACCCTA CACACGACCA CGACAACCAA ACCAGAACCC AGACCACCCT 5160  
GGGCCACCAG TTCCAGACT CGGCCATCAC CCCGCAGAAA GGAAAGGCCA CAACCTGCGC 5220  
ACCCAGCCC CGATCCGGCG GGCAGCCACC CAACCCTAAC CAGCACCCAA GAGCGATCCC 5280  
CGAAGGACCC CCGAACCGCA AAGGACATCA GTATCCCACA GCCTCTCAA GTCCCCGGT 5340  
CTCTTCTCT TCTCGAAGGG ACTAAAAGAT CAATCCACCA CATCCGACGA CACTCAACTC 5400  
CCCGTCCCTA AAGGAGACAC CGGGAATCCC GGAATTAAGA CTCATCCAAT GTCCATCATG 5460  
GGTCTCAAGG TGAACGTCTC TGCCATATTC ATGGCAGTAC TGTTAAGTCT CCAAACACCC 5520  
ACCGGTCAAA TCCATTGGGG CAATCTCTCT AAGATAGGGG TGGTAGGAAT AGGAAGTGCA 5580  
AGCTACAAAG TTATGACTCG TTCCAGCCAT CAATCATTAG TCATAAAATT AATGCCCAAT 5640  
ATAACTCTCC TCAATAACTG CACGAGGGTA GAGATTGCAG AATACAGGAG ACTACTGAGA 5700  
ACAGTTTGG AACCAATTAG AGATGCACTT AATGCAATGA CCCAGAATAT AAGACCGTTT 5760  
CAGAGTGTAG CTTCAAGTAG GAGACACAAG AGATTGCGAG GAGTAGTCCT GGCAGGTGCG 5820  
GCCCTAGGCG TTGCCACAGC TGCTCAGATA ACAGCCGGCA TTGCACTTCA CCAGTCCATG 5880  
CTGAAGTCTC AAGCCATCGA CAATCTGAGG GCAAGTCTGG AAAGTACTAA TCAGGCAATT 5940  
GAGGCAATCA GACAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC 6000  
ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTTAAT CGGCCAGAAG 6060  
CTCGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CATTATTTGG CCCTAGCTTA 6120  
CGGGACCCCA TATCTGCGGA GATATCTATC CAGGCTTTGA GCTATGCGCT CGGAGGAGAT 6180  
ATCAATAAGG TGTTAGAAAA GCTCGGATAT AGTGGAGGTG ATTTACTGGG CATCTTAGAG 6240  
AGCAGAGGAA TAAAGGCCCG GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTCCTC 6300  
AGTATAGCCT ACCCGACGCT GTCCGAGATC AAGGGGGTGA TTGTCCACCG GCTAGAGGGG 6360  
GTCTCGTACA ACATAGGCTC TCAAGAGTGG TATACGACTG TGCCCAAGTA TGTTGCAACC 6420  
CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT 6480

SUBSTITUTE SHEET (RULE 26)

- 110 -

GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG 6540  
TCCACCAAGT CCTGTGCTCG TACACTCGTA TCTGGGTCTT TTGGGAACCG GTTCATTTTG 6600  
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA 6660  
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCCG 6720  
GTAGTCGAGG TGAACGGCGT GACCATCCAA GTCGGGAGCA GGAGGTATCC AGACGCTGTG 6780  
TACTTGACCA GAATTGACCT CGGTCCTCCC ATATCATTGG AGAGGTTGGA CGTAGGGACA 6840  
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC 6900  
CAGATATTGA GGAGTATGAA AGGTTTGTCT AGCACTAGCA TAGTCTACAT CCTGATTGCA 6960  
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT 7020  
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACAGGA 7080  
ACATCGAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGGAAC ACAAATGTCC 7140  
CACAACTCTC CTCTTCGTCA TCAAGCAACC ACCGCATCCA GCATCAAGCC CACCTGAAAT 7200  
TATCTCCGGC TCCCCTTTGG CCGAACAATA TCGGTAGTTA ATTAAAACTT AGGGTGCAAG 7260  
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA 7320  
TCCCAAGGGA AGTAGGATAG TTATCAACAG AGAACACCTT ATGATTGATA GACCTTATGT 7380  
TTTGCTGGCT GTTCTGTTCG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CAATTGCAGG 7440  
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA 7500  
TCTAGATGTA ACTAACTCAA TTGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA 7560  
AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTTACTGACC TAGTGAAATT 7620  
CATCTCTGAC AAGATTAAAT TCCTTAACCC GGATAGGGAG TACGACTTCA GAGATCTCAC 7680  
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT 7740  
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC 7800  
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA 7860  
ATTCTCAAAC ATGTCGCTGT CCCTGTTGGA CTGTGATTTA AGTCGAGGTT ACAATGTGTC 7920  
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTACGGGGGA ACTTACCTAG TGGAAAAGCC 7980  
TAATCTGAGC AGCAAAGGGT CAGAGTTGTC ACAACTGAGC ATGTACCGAG TGTGTTGAAGT 8040

SUBSTITUTE SHEET (RULE 26)

- 111 -

AGGTGTTATC AGAAATCCGG GTTGGGGGCG TCCGGTGTTT CATATGACAA ACTATTTTGA	8100
GCAACCAGTC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAACCTCGC	8160
AGCCCTTTGT CACGGGGGAG ATTCTATCAC AATTCCCTAT CAGGGATCAG GGAAAGGTGT	8220
CAGCTTTCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CCCCTTCTCA ACGGATGACC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTAT CCCGACAACA AGAACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAGC AGGCGTGTA GGTAAAATC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGAGTCTTGT CTGTTGATCT	8520
GAGTCTAACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA	8580
CGTTCAGGG ATGGACCTAT ACAAGTCCAA CCACAACAAT GAGTATTGGC TGACTATCCC	8640
GCCAAATGAAG AACCTAGCCC TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTAGTCCC AACCTCTTCA CTGTCCCAAT TAAGGAAGCA GGCGAAGACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAACCT AGTTCCAATC TGGTGATCCT	8820
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCCAGGG TTGAACATGC	8880
TGTGGTTTAT TACGTTTACA GCCCAAGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT	8940
GCCTATAAAG GGGATCCCCA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC	9060
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATAGCAG	9120
ATAGGGCTGC CAGTGAACCA ATCACATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCCG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ACAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACC GCCG TAAAAACCG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCACT	9480
CTCATATTCC ATATCCAAAC TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540
CGAGGAAGAT CCGTGAACCT CTCAAAAAGG GAAATTCGCT GTACTCTAAA GTCAGTAATA	9600

SUBSTITUTE SHEET (RULE 26)

- 112 -

AGGTTTTCCA ATGCTTGAGG GACACTAATT CACGGCTTGG TCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAACCTTG GAGTTTACAT GCACAGCTCC CAATGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TAACATTTGA GCTGGTTTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CTGCTATGAC CATTGATGCT AGATATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AATTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTTCACTTGCT TACCTGCAGC TGAGGGATAT AACGGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCACGAG TTAGTTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA	10260
TACACCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC CTCAGGTGAA GGATTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT	10560
TTGCTGGAGT GAAATTTGGC TGCTTCATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT	10620
ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG	10680
AGTTCCTGCG TTACGACCCC CCCAAGGGAA CCGGGTCAGG GAGGCTTGTA GATGTTTTCC	10740
TTAATGATTC GAGCTTTGAC CCATATGATA TGATAATGTA TGTTGTAAGT GGAGCTTACC	10800
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG	10860
GTAGACTTTT TGCTAAAATG ACTTACAAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC	10920
TAATCTCAAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCAGC	10980
ATTTGACTAA GGCCTCCAC ACTCTGGCTG TCTCAGGAGT CCCTAAAGAT CTCAAAGAAA	11040
GTCACAGAGG GGGGCCAGTC CTAAAAACCT ACTCCCGAAG CCCAGCCCAC ACAAATACCA	11100
GGAACGTGAG GGCAGCAAAA GGGTTTATAG GGTTCCTCA GATAATTCCG CAGGACCAAG	11160

SUBSTITUTE SHEET (RULE 26)



- 113 -

ACACTAATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATTT ATCACAACCTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTA CCCTCATTTT TTCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCCTGTA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCCAT ATCCCGTTAT 11400  
GCAAAGTCCC CAATGACCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATT CCTATTTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCATTA GTGCAAGGGG ACAATCAGAC CATAGCTGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCTTAC AACCTTAAGA AATGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ACATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTTTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCTCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CCTTGCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAG ATTCTGATCT 11940  
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCAGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCATCACTGA TGCCTGAAGA GACCCTCCAT CAAGTAATGA 12180  
CACAGCAACC GGGGGACTCT TCATTCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240  
TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTAA 12300  
TCCACAGTCC AAACCCAATG TTAAAGGGAT TATTCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC AGCATTCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCACA GGGGCAAGAG AGTCTATTGC AGGCATGCTA GATACCACAA 12480  
AAGGCCTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGAATATGAA CAATTCTAGAG CAGGGATGGT GCTATTAAACA GGAAGAAAGA 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GCGAGAGCC CTAAGAAGCC 12660  
ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC 12720

SUBSTITUTE SHEET (RULE 26)

- 114 -

TAGAATCTAT GCGAGGCCAC CTTATTCCGGC GTCATGAGAC ATGTGTCATC TCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTGTGTC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTCGTAAGA GCCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA 12960  
CAGTGTA CTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG 13020  
CAAGGCAAAG GGCTAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCACATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTC CG AGTGGCAAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT 13200  
CAGATAAGAA GGTTGATACT AACTTTATAT ACCAACAAGG AATGCTCCTA GGGTTGGGCG 13260  
TTT TAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TGC GTGATCC CAATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CTCGCAAGCT AGAGCTGAGG GCAGAGCTGT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTAG 13500  
AATTTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTCT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CCTGGTAACA AAATTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTTATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740  
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTTGTCATC GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTC AATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCACTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGACAA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTTACA TTTCTTTTGT GTGAAAGTGA CGAGGATGTA GTACCGGACA 14040  
GATTCGACAA CATCCAGGCA AAACACTTGT GTGTTCTGGC AGATTTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCAATTGCA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC 14160  
ATATCAAGGC GGAGGCTAGG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTA CT CATGC TCTCTGACTT ATCTTCGGCG AGGATCGATC AAACAGATAA 14280

SUBSTITUTE SHEET (RULE 26)

- 115 -

GATTGAGAGT TGATCCAGGA TTCATTTTCG ACGCCCTCGC TGAGGTAAAT GTCAGTCAGC 14340  
CAAAGATCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGATTTCAGA CCCCCACACG 14400  
ATGATGTTGC AAAATTGCTC AAAGATATCA ATACAAGCAA GCACAATCTT CCCATTTCTG 14460  
GGGGCAATCT CGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTG AACTCATCTG 14520  
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAAG 14580  
ACGGCTTATT CTTGGGTGAG GGATCGGGTT CTATGTTGAT CACTTATAAG GAGATACTTA 14640  
AACTAAACAA GTGCTTCTAT AATAGTGGGG TCTCTGCCAA TTCTAGATCT GGTCAAAGGG 14700  
AATTAGCACC CTATCCCTCC GAAGTTGGCC TTGTGGAACA CAGAATGGGA GTAGGTAATA 14760  
TTGTCAAGGT GCTCTTTAAC GGGAGGCCCG AAGTCACATG GGTAGGCAGT GTAGATTGCT 14820  
TCAATTACAT AGTTAGTAAT ATCCCTACCT CTAGTGTGGG GTTTATCCAT TCAGATATAG 14880  
AGACCTTACC TAACAAAGAT ACTATAGAGA AGCTAGAGGA ATTGGCAGCC ATCTTATCGA 14940  
TGGCTCTGCT CCTGGGCAAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTACAGC 15000  
GGGATTTTGT TCAGGGATTT ATAAGTTATG TAGGGTCTCA TTATAGAGAA GTGAACCTTG 15060  
TATACCCAG ATACAGCAAC TTCATATCTA CTGAATCTTA TTTGGTTATG ACAGATCTCA 15120  
AGGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGC 15180  
GGACTTCACC TGGACTTATA GGTACATCC TATCCATTAA GCAACTAAGC TGCATACAAG 15240  
CAATTGTGGG AGACGCAGTT AGTAGAGGTG ATATCAATCC TACTCTGAAA AAACCTACAC 15300  
CTATAGAGCA GGTGCTGATC AATTGCGGGT TGGCAATTAA CGGACCTAAA CTGTGCAAAG 15360  
AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAATTCT ATACTCATCC 15420  
TCTACAGGGA GTTGGCAAGA TTCAAGGACA ACCAAAGAAG TCAACAAGGG ATGTTCCACG 15480  
CTTACCCCGT ATTGGTAAGT AGCAGGCAAC GAGAACTTAT ATCTAGAATC ACTCGCAAAT 15540  
TTTGGGGGCA CATTCTTCTT TACTCCGGGA ACAGAAAGTT GATAAATAAG TTTATCCAGA 15600  
ATCTCAAGTC CGGTTATCTG ATACTAGACT TACACCAGAA TATCTTCGTT AAGAATCTAT 15660  
CCAAGTCAGA GAAACAGATT ATTATGACGG GGGGTTTGAA ACGTGAGTGG GTTTTTAAGG 15720  
TAACAGTCAA GGAGACCAAG GAATGGTATA AGTTAGTCGG ATACAGTGCC CTGATTAAGG 15780  
ACTAATTGGT TGAACCTCCG AACCCCTAATC CTGCCCCAGG TGGTTAGGCA TTATTTGTAA 15840

SUBSTITUTE SHEET (RULE 26)

- 116 -

TATATTAAAG AAAACTTTGA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT

15894

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu
1           5           10           15

Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala
20          25          30

Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn
35          40          45

Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn
50          55          60

Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro
65          70          75          80

Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn
85          90          95

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys
100         105         110

Gly Asn Ser Leu Tyr Ser Lys Val Ser Asn Lys Val Phe Gln Cys Leu
115         120         125

Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp
130         135         140

Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln
145         150         155         160

Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg
165         170         175

Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr
180         185         190

```

SUBSTITUTE SHEET (RULE 26)

- 117 -

Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
 195 200 205

Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
 210 215 220

Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
 225 230 235 240

Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly  
 245 250 255

Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
 260 265 270

Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu  
 275 280 285

Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe  
 290 295 300

Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly  
 305 310 315 320

Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Val Glu Ala Leu Asp Tyr  
 325 330 335

Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe  
 340 345 350

Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu  
 355 360 365

Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr  
 370 375 380

Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr  
 385 390 395 400

Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His  
 405 410 415

Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr  
 420 425 430

His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe  
 435 440 445

Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu  
 450 455 460

SUBSTITUTE SHEET (RULE 26)

- 118 -

Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr  
 465 470 475 480  
 Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg  
 485 490 495  
 Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp  
 500 505 510  
 Met Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe  
 515 520 525  
 Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg  
 530 535 540  
 Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala  
 545 550 555 560  
 Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly  
 565 570 575  
 Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala  
 580 585 590  
 Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro  
 595 600 605  
 Val Leu Lys Thr Tyr Ser Arg Ser Pro Ala His Thr Asn Thr Arg Asn  
 610 615 620  
 Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Ile Ile Arg Gln  
 625 630 635 640  
 Asp Gln Asp Thr Asn His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val  
 645 650 655  
 Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670  
 Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly  
 675 680 685  
 Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val  
 690 695 700  
 Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile  
 705 710 715 720  
 Pro Leu Cys Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met  
 725 730 735  
 Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile

SUBSTITUTE SHEET (RULE 26)

- 119 -

740					745					750					
Pro	Tyr	Leu	Tyr	Leu	Ala	Ala	Tyr	Glu	Ser	Gly	Val	Arg	Ile	Ala	Ser
		755					760					765			
Leu	Val	Gln	Gly	Asp	Asn	Gln	Thr	Ile	Ala	Val	Thr	Lys	Arg	Val	Pro
		770					775					780			
Ser	Thr	Trp	Pro	Tyr	Asn	Leu	Lys	Lys	Trp	Glu	Ala	Ala	Arg	Val	Thr
							790					795			800
Arg	Asp	Tyr	Phe	Val	Ile	Leu	Arg	Gln	Arg	Leu	His	Asp	Ile	Gly	His
				805					810					815	
His	Leu	Lys	Ala	Asn	Glu	Thr	Ile	Val	Ser	Ser	His	Phe	Phe	Val	Tyr
				820					825					830	
Ser	Lys	Gly	Ile	Tyr	Tyr	Asp	Gly	Leu	Leu	Val	Ser	Gln	Ser	Leu	Lys
			835				840					845			
Ser	Ile	Ala	Arg	Cys	Val	Phe	Trp	Ser	Glu	Thr	Ile	Val	Asp	Glu	Thr
			850				855					860			
Arg	Ala	Ala	Cys	Ser	Asn	Ile	Ala	Thr	Thr	Met	Ala	Lys	Ser	Ile	Glu
				870								875			880
Arg	Gly	Tyr	Asp	Arg	Tyr	Leu	Ala	Tyr	Ser	Leu	Asn	Val	Leu	Lys	Val
				885					890					895	
Ile	Gln	Gln	Ile	Leu	Ile	Ser	Leu	Gly	Phe	Thr	Ile	Asn	Ser	Thr	Met
			900					905					910		
Thr	Gln	Asp	Val	Val	Ile	Pro	Leu	Leu	Thr	Asn	Asn	Asp	Leu	Leu	Ile
			915				920					925			
Arg	Met	Ala	Leu	Leu	Pro	Ala	Pro	Ile	Gly	Gly	Met	Asn	Tyr	Leu	Asn
			930				935					940			
Met	Ser	Arg	Leu	Phe	Val	Arg	Asn	Ile	Gly	Asp	Pro	Val	Thr	Ser	Ser
				950								955			960
Ile	Ala	Asp	Leu	Lys	Arg	Met	Ile	Leu	Ala	Ser	Leu	Met	Pro	Glu	Glu
				965					970					975	
Thr	Leu	His	Gln	Val	Met	Thr	Gln	Gln	Pro	Gly	Asp	Ser	Ser	Phe	Leu
				980					985					990	
Asp	Trp	Ala	Ser	Asp	Pro	Tyr	Ser	Ala	Asn	Leu	Val	Cys	Val	Gln	Ser
			995				1000					1005			
Ile	Thr	Arg	Leu	Leu	Lys	Asn	Ile	Thr	Ala	Arg	Phe	Val	Leu	Ile	His
			1010				1015					1020			

SUBSTITUTE SHEET (RULE 26)

- 120 -

Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
 1025 1030 1035 1040  
 Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
 1045 1050 1055  
 Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
 1060 1065 1070  
 Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
 1075 1080 1085  
 Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
 1090 1095 1100  
 Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly  
 1105 1110 1115 1120  
 Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu  
 1125 1130 1135  
 Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg  
 1140 1145 1150  
 Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly  
 1155 1160 1165  
 His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser  
 1170 1175 1180  
 Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp  
 1185 1190 1195 1200  
 Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr  
 1205 1210 1215  
 Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser  
 1220 1225 1230  
 Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala  
 1235 1240 1245  
 Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg  
 1250 1255 1260  
 Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile  
 1265 1270 1275 1280  
 Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln  
 1285 1290 1295

SUBSTITUTE SHEET (RULE 26)



- 121 -

Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr  
 1300 1305 1310  
 Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp  
 1315 1320 1325  
 Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu  
 1330 1335 1340  
 Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val  
 1345 1350 1355 1360  
 Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp  
 1365 1370 1375  
 His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu  
 1380 1385 1390  
 Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp  
 1395 1400 1405  
 Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe  
 1410 1415 1420  
 Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr  
 1425 1430 1435 1440  
 Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met  
 1445 1450 1455  
 Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile  
 1460 1465 1470  
 Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly  
 1475 1480 1485  
 Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro  
 1490 1495 1500  
 Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg  
 1505 1510 1515 1520  
 Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro  
 1525 1530 1535  
 Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His  
 1540 1545 1550  
 Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met  
 1555 1560 1565  
 Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu

SUBSTITUTE SHEET (RULE 26)

- 122 -

1570	1575	1580
Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val 1585	1590	1595 1600
Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala 1605	1610	1615
Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg 1620	1625	1630
Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala 1635	1640	1645
Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val 1650	1655	1660
Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys 1665	1670	1675 1680
Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala 1685	1690	1695
Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn 1700	1705	1710
Met Ser Ile Lys Asp Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu 1715	1720	1725
Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly 1730	1735	1740
Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn 1745	1750	1755 1760
Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg 1765	1770	1775
Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly 1780	1785	1790
Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe 1795	1800	1805
Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu 1810	1815	1820
Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val 1825	1830	1835 1840
Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp 1845	1850	1855

SUBSTITUTE SHEET (RULE 26)

- 123 -

Val Gly Ser Val Asp Cys Phe Asn Tyr Ile Val Ser Asn Ile Pro Thr  
 1860 1865 1870  
 Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys  
 1875 1880 1885  
 Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
 1890 1895 1900  
 Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
 1905 1910 1915 1920  
 Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His  
 1925 1930 1935  
 Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser  
 1940 1945 1950  
 Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met  
 1955 1960 1965  
 Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr  
 1970 1975 1980  
 Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys  
 1985 1990 1995 2000  
 Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro  
 2005 2010 2015  
 Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly  
 2020 2025 2030  
 Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp  
 2035 2040 2045  
 Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr  
 2050 2055 2060  
 Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met  
 2065 2070 2075 2080  
 Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile  
 2085 2090 2095  
 Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly  
 2100 2105 2110  
 Asn Arg Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr  
 2115 2120 2125

SUBSTITUTE SHEET (RULE 26)

- 124 -

Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
 2130 2135 2140

Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
 2145 2150 2155 2160

Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
 2165 2170 2175

Tyr Ser Ala Leu Ile Lys Asp  
 2180

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ACCAAACAAA GTTGGGTAAG GATAGATCAA TCAATGATCA TATTCTAGTA CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCCTCAA	240
TTACCACTCG ATCCAGACTA CTGGACCGGT TGGTCAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTGTTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATCAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAAGTAG TAGTGATCAA TCCAGGTCCG	540
GATGGTTCGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATTCTA GCCCAAATTT GGGTCTTGCT CGCGAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTCAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780

SUBSTITUTE SHEET (RULE 26)

- 125 -

AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGGACACCCG	840
GGAACAAACC AAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTA ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GACTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAT CTTTACCAGC	1020
AAATGGGAGA AACTGCACCC TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAATTCA	1080
GTGCAGGATC ATACCCCTG CTCTGGAGCT ATGCCATGGG AGTAGGGGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTGGTTCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTGAGGAGG TCAGCTGGGA AAGTCAGTTC CACATTAGCA TCTGAACTCG	1260
GTATCACTGC TGAGGATGCA AGGCTTGTTT CAGAGATTGC AATGCACACT ACTGAGGACA	1320
GGACCAGTAG AGCGGTTGGA CCCAGACAAG CCCAAGTGTC ATTTCTACAC GGTGATCAAA	1380
GTGAGAATGA GCTACCAGGA TTGGGGGGCA AGGAAGATAG GAGGGTCAAA CAGAGTCGGG	1440
GAGAAGCCAG GGAGAGCTAC AGAGAAACCG GGTCTAGCAG AGCAAGCGAT GCGAGAGCTG	1500
CCCATCTTCC AACCAGCGCA CCCCTAGACA TTGACACTGC ATCGGAGTCA GGCCAAGATC	1560
CGCAGGACAG TCGACGGTCA GCTGACGCCC TGCTCAGGCT GCAAGCCATG GCAGGAATCT	1620
TGGAAGAACA AGGCTCAGAC ACGGACACCC CTAGGGTGTA CAATGACAGA GATCTTCTAG	1680
ACTAGGTGCG AGAGGCCGAG GACCAGAACA ACATCCGCCT ACCCTCCATC ATTGTTATAA	1740
AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCAACCA ACCATCCACT CCTACGACTG	1800
GGGCCGATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGCT	1860
CTCAAGGCCG AGCCCATCGG CTCCTGGCC GTCGAGGAAG CCATGGCAGC ATGGTCACAA	1920
ATATCAGACA ACCCAGGACA GGACCGAACC ACCCGCAAGG AAGAGGAGGC AGGCAGTTCG	1980
GGTCTCAGCA AACCATGCCT CTCAGCAATT GGATCAACTG AAGGCAGTGC ACCTCGCATC	2040
TGCGGTTCAGG GATCTGGAGA GAGCGATGAC AACGCTGAAA CTTTGGAAT CCCCTCAAGA	2100
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTTATCATG TTTATGATCA CAGCGGTGAA	2160
GCGGTTAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT	2220
AGCACCTCT CAGGAGGAGA CGATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT	2280
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTGAGG	2340

SUBSTITUTE SHEET (RULE 26)

- 126 -

GCTTCTGATG TTGAACTGC AGAAGGAGGG GAGATCCACG AGCTCCTGAG ACTCCAATCT	2400
AGAGGCAACA ACTTCCCGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC CCCGAACCCC	2460
GGTAGGGCCA GCACTTCCGA GACACCCATT AAAAAGGGGA CAGACGCGAG ATTAGCCTCA	2520
TTTGGAGCGG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA	2580
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GTGGGGAATG TCCCCGAGTG TGTGAGCAAT	2640
CCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCAG	2700
AATAATGAAG AAGGGGGAGA TTATTATGAT GATGAGCTGT TCTCCGATGT CCAAGACATC	2760
AAAACAGCCT TGGCCAAAT ACACGAGGAT AATCAGAAGA TAATCTCCAA GCTAGAATCA	2820
CTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAAAAGC AGATCAACAG GCAAAATATC	2880
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG	2940
AAGGATCCCA ACGACCCAC TGCAGATGTC GAACTCAATC CCGACCTGAA ACCCATCATA	3000
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCGTTGC CAGCCGACAA	3060
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTCAA	3120
CTAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCTGA CACCGGCCCC	3180
GCATCACGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG	3240
CGTTACCTGA TGACTCTCCT TGATGATATC AAAGGAGCCA ACGATCTTGC CAAGTTCCAC	3300
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCTCATG	3360
CCAATCGACC TAATTAGTAC AGCCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT	3420
GCCTCCCAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA	3480
AAGGGTCGAT CGCTCCGATA CAACCTACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG	3540
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTACG TACATGTTTC	3600
TGCTGGGGGT TGTTGAGGAC AGCGATCCCC TAGGGCCTCC AATCGGGCGA GCATTTGGGT	3660
CCCTGCCCTT AGGTGTTGGT AGATCCACAG CAAAACCCGA AGAACTCCTC AAAGAGGCCA	3720
CTGAGCTTGA CATAGTCGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA	3780
ACAACACCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT	3840
TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTGGATACC CCGCAGAGGT	3900

SUBSTITUTE SHEET (RULE 26)

- 127 -

TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTTCCTA	3960
GAAGAATGCT AGAATTCAGA TCGGTCAATG CAGTGGCTTT CAACCTGCTG GTGACCCCTTA	4020
GGATTGACAA AGCGATTGGC CCTGGGAAGA TCATCGATAA TGCAGAGCAA CTCCTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCTG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCACCTGGT GGGATAGGGG	4200
GCACCAGTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACCTG	4260
GGTTCAAAAA GACCTTATGT TACCCACTGA TGGATATCAA TGAAGACCTT AATCGATTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCCC	4380
AAGAATTCCG CATTTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC	4440
TGTAGACCGT AGTGCCCGC AATACCCGAA AACGACCCCC CTCATAATGA CAGCCAGAAG	4500
GCCCGGACAA AAAAGCCCCC TCCAAAAGAC TCCACGGACC AAGTGAGAGG CCAGCCAGCA	4560
GCTGACGGCA AGCGTGAACA CCAGGCGGCC TGGGCACAGA ACAGCCCCGA CACAAGGCAA	4620
CCACCAGCCA TCCCAATCTG CGTCCTCCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGT	4680
CGCCCCGAC CCAGACCACC AACC GCATCC CCACAGCCCC CGGGAAAGAG ACCCCAGCA	4740
ACTGGAAGGC CCCTCCCCCT TTCCCTCAAC GCAAGAACTC CACAACCGAA CCGCACAAGC	4800
GATCGAGGTG ACCCAACCGC AGGCATCCGA CTCCTTAGAC AGATCCTCTC CCCCCGGCAA	4860
ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCGAC AGAACCAGA CCCC GGCCCA	4920
CGGCGCCGCG CCCCCACCTC CCGACAACCA GAGGGAGCCC CCAACCAATC CCGCCGGCTC	4980
CCCCGGTGCC CACAGGCAGG CACACCAACC CTCGAACAGA CCCAGCACCC AGCCATCGAC	5040
AATTCAAGAC GGGGGGCCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCACCGC	5100
GAGGAAGCCC ACCCACCCTA CACACGACCA CAGGAACCGA ACCAGAATCC AGACCACCCT	5160
GGGCCACCAG TTCCAGACT CGGCCATCAC CCCGCAGAAA GGAAAGGCCA CAACCCGCGC	5220
ACCCCTGCCC TGATCCGGTG GCGGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC	5280
CGAAGGGCCC CCGAACCGCA AAAGACATCA GTATCCACA GCCTCTCAA GTCCCCGGT	5340
CTCCCCCTCT TCTCGAAGGG ACCAAAAGAT CAATCCACCA CACCCGACGA CACTCAATTC	5400
CCCACCCCTA AAGGAGACAC CGGGAATCCC AGAATCAAGA CTCATCCAAT GTCCATCATG	5460

SUBSTITUTE SHEET (RULE 26)

- 128 -

GGTCTCAAGG TGAACGTCTC TGCCATATTC ATGGCAGTAC TGTAACTCT CCAAACACCC	5520
ACCGGTCAAA TCCATTGGGG CAATCTCTCT AAGATAGGGG TGGTAGGGAT AGGAAGTGCA	5580
AGCTACAAAG TTATGACTCG TTCCAGCCAT CAATCATTAG TCATAAAATT AATGCCCAAT	5640
ATAACTCTCC TCAATAACTG CACGAGGGTA GAGATTGCAG AATACAGGAG ACTACTGAGA	5700
ACAGTTTTTG AACCAATTAG AGATGCACTT AATGCAATGA CCCAGAATAT AAGACCGGTT	5760
CAGAGTGTAG CTTCAAGTAG GAGACACAAG AGATTTGCTG GAGTTGTCCT GCGGGGTGCG	5820
GCCCTAGGCG TTGCCACAGC TGCTCAGATA ACAGCCGGA TTGCACTTCA CCAGTCCATG	5880
TTGAACTCTC AAGCCATCGA CAATCTGAGA GCGAGCCTGG AAACACTAA TCAGGCAATT	5940
GAGGCAATCA GACAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC	6000
ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTAAAT CGGCCAGAAG	6060
CTAGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CACTATTTGG CCCCAGCTTA	6120
CGGGACCCCA TATCTGCGGA GATATCTATC CAGGCTTTGA GCTATGCGCT TGGAGGAGAT	6180
ATCAATAAGG TGTTAGAAAA GCTCGGATAC AGTGGAGGTG ATTTACTGGG CATCTTAGAG	6240
AGCAGAGGAA TAAAGGCCCG GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTACTC	6300
AGTATAGCCT ATCCGACGCT GTCCGAGATT AAGGGGGTGA TTGTCCACCG GCTAGAAGGG	6360
GTCTCGTACA ACATAGGCTC TCAAGAGTGG TATACCACTG TGCCCAAGTA TGTTGCAACC	6420
CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT	6480
GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TCCACCAAGT CCTGTGCTCG TACACTTGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ATCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAGGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCCG	6720
GTGGTCGAGG TGAACGGCGT GACCATCCAA GTCGGGAGCA GCGGGTATCC GGACGCTGTG	6780
TACTTGCACA GAATTGACCT CGGTCTCTCC ATATCATTGG AGAGGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTAGCA TAGTTTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020

SUBSTITUTE SHEET (RULE 26)



- 129 -

AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACAGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCATCCA GCATCGAGCC CACCTGAAAT	7200
TGTCTCCGGA TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATTAAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ACAACCCCCA	7320
TCCTAGGGGA AGTAGGATAG TTATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTATTCTG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATAAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500
TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
GATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTCACCGACC TAGTGAAATT	7620
CATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGTATC AATCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAACTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGG CCAGGGTAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCCTGTTGGA CTTGTATTTA AATCGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTACGGGGGA ACTTACCTAG TGGAAAAGCC	7980
TAATCTGAGC AGTAAAGGGT CAGAGTTGTC ACAACTGAGC ATGCACCGAG TGTGTTGAAGT	8040
AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTC CATATGACAA ACTATTTTGA	8100
GCAACCAGTC AGTAATGATT TCAGCAACTG CATGGTGGCT TTGGGGGAGC TCAAATTCGC	8160
AGCCCTTTGT CACAGGGAAG ATTCTATCAC AATTCCCTAT CAGGGATCAG GGAAAGGTGT	8220
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CCCCCTATCA ACGGATGATC CAGTGATAGA CAGGCTCTAC CTCTCATCTC ACAGAGGCGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGGACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAGC AGGCGTGTA GGGTAAATC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTAATCT	8520
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCAGGA TTCGGGCCAT TGATCACACA	8580

SUBSTITUTE SHEET (RULE 26)

- 130 -

CGGTTTCAGGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGACTATCCC	8640
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTTAGTCCC TACCTCTTCA CTGTTCCAAT TAAGGAAGCA GCGGAGGACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAATC AGTTCCAATC TGGTGATTCT	8820
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTATGAT ACTTCCAGAG TTGAACATGC	8880
TGTGGTTTAT TACGTTTACA GCCCAAGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT	8940
GCCTATAAGG GGGGTCCCCA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGATATA TCACTCACTC	9060
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACTCGG GAAGATGGAA CCAACCGCAG	9120
ATAGGGCTGC CAGTGAACCA ATCACATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCTATCCTG GAGTATGCTC GAGTCCCTCA CGCATACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACCGCC TAAAAACGG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGACCCACT	9480
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540
CAAGGAAGAT CCGTGAGCTC CTCAAAAAGG GAAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCCTGAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAATT ATTAAGTTGG GAGTTTACAT GCACAGCTCC CAATGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CAGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACGTTTGA ACTGGTCTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC CATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTACCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTCATTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG	10140

SUBSTITUTE SHEET (RULE 26)

- 131 -

CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG 10200  
ATGAAGGTAC TTATCATGAG TTAATTGAAG CCCTAGATTA CATTTTCATA ACTGATGACA 10260  
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG 10320  
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG 10380  
AGACTCTGAT GAAAGGTCAT GCCATATTCT GTGGAATCAT AATCAACGGC TATCGTGACA 10440  
GGCACGGAGG CAGTTGGCCA CCCCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA 10500  
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT 10560  
TTGCTGGAGT GAAATTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT 10620  
ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG 10680  
AGTTCCTGCG TTACGACCCT CCCAAAGGAA CTGGGTCACG GAGGCTTGTA AATGTTTTCC 10740  
TTAATGATTC GAGCTTTGAC CCATATGACA TGATAATGTA TGTGTAAAGT GGAGCTTACC 10800  
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG 10860  
GTAGACTTTT TGCTAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC 10920  
TAATCTCAA CGGGATTGGC AATTATTTTA AGGACAATGG GATGGCCAAG GACGAGCACG 10980  
ATTTGACTAA GGCCTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA 11040  
GTCACAGGGG GGGGCCAGTC TTA AAAACCC ACTCCGAAG CCCAGTCCAC ACAAGTACCA 11100  
AGAACGTGAG AGCAGCAAAA GGGTTTATAG GATTCCCTCA TGTAATTCGG CAGGACCAAG 11160  
ACACTGATCA TCCGGAGAAT ATGGAGGCTT ACGAGACAGT CAGTGCATTT ATCAGGACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTATTT GCACAAAGGC 11280  
TAAATGAGAT TTACGGATTA CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAAACCT 11340  
CTGTCCTCTA TGTAAGTGAC CCTCATTGCC CCCCTGACCT TGACGCCCAT GTCCCGTTAT 11400  
GCAAAGTCCC CAATGACCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAAGT GTGGACCATC AGCACCATTG CCTATTTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCTTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ACATAGGCCA TCACCTCAAG GCAAATGAGA 11700

SUBSTITUTE SHEET (RULE 26)

- 132 -

CAATTGTCTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCTCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CCTTGCAAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATCCTGATCT 11940  
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGATCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATCGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC TCATCACTAA TGCCTGAAGA GACCCTTCAT CAAGTAATGA 12180  
CACAAACAACC GGGGGACTCT TCATTCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240  
TTGTATGCGT CCAGAGCATC ACTAGACTCC TCAAGACAT AACTGCAAGG TTTGTCCTGA 12300  
TCCATAGTCC AAACCCAATG TTAAAAGGGT TATTCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC GGCATTCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCA CA GGGGCAAGAG AGTCTATTGC AGGCATGCTA GATACCACAA 12480  
AAGGCCTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGACTATGAA CAATTTAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GGCTAGAGCC CTAAGAAGCC 12660  
ATATGTGGGC AAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC 12720  
TAGAATCTAT GCGAGGCCAC CTTATTCGGC GCCATGAGAC ATGTGTCATC TCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTGTGTC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTGTAAGA GCCCCAAGTC GATCCTTGC ATCTGCTGTT AGAATAGCAA 12960  
CAGTGACTC ATGGGCTTAT GGTGATGATG ATAGCTCTG GAACGAAGCC TGGTTGTTGG 13020  
CAAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCGCATAGG TTGAGGGATC GTACCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTCG AGTGGCAAGG TATACCACAA TCTCAACGA CAATCTCTCA TTTGTCATAT 13200  
CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAGGG AATGCTTCTA GGGTTGGGTG 13260

SUBSTITUTE SHEET (RULE 26)

- 133 -

TTTGTAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TCGGTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CCCGCAAGCT AGAGCTTAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG 13500  
AATTTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTTATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCAGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740  
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTGTCTTCT GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTACACA 13920  
CAACTGTGTG CAACATGATT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTTCTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA 14040  
GATTCGACAA TATCCAGGCA AAACACTTGT GTGTTCTAGC AGATTTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCAATTCGA GGTCTACGAC CTGTAGAGAA ATGTGCAGTT CTAACCGATC 14160  
ATATCAAGGC AGAGGCTAGG TTATCTCCAG CAGGGTCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTAATCATGC TCTCTGACTT ATCTCCGGCG AGGATCGATC AAACAGATAA 14280  
GATTGAGAGT TGATCCAGGA TTCATTTTGT ACGCCCTCGC TGAGGTAAAT GTCAGTCAGC 14340  
CAAAGATCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGATTTGAGA CCTCCACAGG 14400  
ATGATGTTGC AAAATTGCTC AAAGATATCA ACACAAGCAA GCACAATCTT CCCATTTGAG 14460  
GGGGTAATCT CGCCAATTAT GAAATCCACG CTTTCCGCAG AATCGGGTTA AACTCATCCG 14520  
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAAG 14580  
ACGGCTTGTT CTTGGGTGAG GGGTCGGGTT CTATGTTGAT CACTTATAAG GAGATACTAA 14640  
AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG 14700  
AATTAGCACC CTATCCCTCC GAAGTTGGTC TTGTCGAACA CAGAATGGGA GTAGGTAATA 14760  
TTGTCAAAGT GCTCTTTAAC GGGAGGCCCG AAGTCACGTG GGTAGGCAGT GTAGATTGCT 14820

SUBSTITUTE SHEET (RULE 26)

- 134 -

TCAATTTCAI AGTCAGTAAT ATCCCTACCT CTAGTGTGGG GTTTATCCAT TCAGATATAG 14880  
AGACCTTACC TAACAAAGAT ACTATAGAGA AGCTAGAGGA ATTAGCAGCC ATCTTATCGA 14940  
TGGCTCTGCT CCTTGGCAAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTCAGCG 15000  
GGGATTTTGT TCAGGGATTT ATAAGTTATG TAGGGTCTTA TTATAGAGAA GTGAACCTTG 15060  
TCTACCCTAG ATACAGCAAC TTCATATCTA CTGAATCTTA TTTAGTCATG ACAGATCTCA 15120  
AAGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGC 15180  
GGACTTCACC TGGACTTATA GGTCAATCC TATCCATTAA GCAACTAAGC TGCATACAAG 15240  
CAATTGTGGG AGACGCAGTT AGTAGAGGTG GTATCAACCC TATTCTGAAG AAACCTACAC 15300  
CTATAGAGCA GGTGCTGATC AATTGCGGGT TGGCAATTAA CGGACCTAAA CTGTGCAAAAG 15360  
AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAACTCT ATACTCATCC 15420  
TCTACAGGGA GTTGGCAAGA TTCAAAGACA ACCAAAGAAG TCAACAAGGG ATGTTCCATG 15480  
CTTACCCCGT ATTGGTAAGT AGCAGGCAAC GAGAACTTAT ATCTAGGATC ACCCGCAAAT 15540  
TTTGGGGGCA TATTCTTCTT TACTCCGGGA ACAGAAAGTT GATAAATCGG TTTATCCAGA 15600  
ATCTCAAGTC CGGTTACCTG ATACTAGACT TACACCAGAA TATCTTCGTT AAGAATCTAT 15660  
CTAAGTCAGA GAAACAGATT ATTATGACGG GGGGTTTAAA ACGTGAGTGG GTTTTAAAGG 15720  
TAACAATCAA GGAGACCAAA GAATGGTATA AGTTAGTCGG ATACAGTGCC CTGATTAAGG 15780  
ATTAATTGGT TGGACTCCGG GACCCTAATC CTGCCCTAGG TAGTTAGGCA TTATTTGCAA 15840  
TATATTAAAG AAAACTTTGA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT 15894

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

- 135 -

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
 1 5 10 15  
 Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
 20 25 30  
 Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
 35 40 45  
 Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
 50 55 60  
 Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
 65 70 75 80  
 Thr His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn  
 85 90 95  
 Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys  
 100 105 110  
 Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu  
 115 120 125  
 Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp  
 130 135 140  
 Ile Lys Glu Lys Ile Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln  
 145 150 155 160  
 Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg  
 165 170 175  
 Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr  
 180 185 190  
 Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
 195 200 205  
 Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
 210 215 220  
 Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
 225 230 235 240  
 Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly  
 245 250 255  
 Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
 260 265 270  
 Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu

SUBSTITUTE SHEET (RULE 26)

- 136 -

275	280	285
Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe 290	295	300
Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly 305	310	315 320
Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr 325	330	335
Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe 340	345	350
Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu 355	360	365
Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr 370	375	380
Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr 385	390	395 400
Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His 405	410	415
Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr 420	425	430
His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe 435	440	445
Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu 450	455	460
Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr 465	470	475 480
Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg 485	490	495
Arg Leu Val Asn Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp 500	505	510
Met Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe 515	520	525
Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg 530	535	540
Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala 545	550	555 560

SUBSTITUTE SHEET (RULE 26)



- 137 -

Glu Asn Leu Ile Ser Asn Gly Ile Gly Asn Tyr Phe Lys Asp Asn Gly  
 565 570 575

Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala  
 580 585 590

Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro  
 595 600 605

Val Leu Lys Thr His Ser Arg Ser Pro Val His Thr Ser Thr Lys Asn  
 610 615 620

Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro His Val Ile Arg Gln  
 625 630 635 640

Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val  
 645 650 655

Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670

Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly  
 675 680 685

Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val  
 690 695 700

Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Val  
 705 710 715 720

Pro Leu Cys Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met  
 725 730 735

Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile  
 740 745 750

Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser  
 755 760 765

Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro  
 770 775 780

Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr  
 785 790 795 800

Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His  
 805 810 815

His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr  
 820 825 830

SUBSTITUTE SHEET (RULE 26)

- 138 -

Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845  
 Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860  
 Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880  
 Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895  
 Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910  
 Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile  
 915 920 925  
 Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940  
 Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960  
 Ile Ala Asp Leu Lys Arg Met Ile Leu Ser Ser Leu Met Pro Glu Glu  
 965 970 975  
 Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990  
 Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005  
 Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His  
 1010 1015 1020  
 Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
 1025 1030 1035 1040  
 Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
 1045 1050 1055  
 Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
 1060 1065 1070  
 Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
 1075 1080 1085  
 Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
 1090 1095 1100  
 Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly

SUBSTITUTE SHEET (RULE 26)

- 139 -

1105	1110	1115	1120
Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu			
1125		1130	1135
Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg			
1140	1145		1150
Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly			
1155	1160		1165
His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser			
1170	1175		1180
Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp			
1185	1190	1195	1200
Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr			
1205	1210		1215
Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser			
1220	1225		1230
Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala			
1235	1240		1245
Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg			
1250	1255		1260
Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile			
1265	1270	1275	1280
Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Thr Thr Gln			
1285	1290		1295
Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr			
1300	1305		1310
Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp			
1315	1320		1325
Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu			
1330	1335		1340
Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val			
1345	1350	1355	1360
Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp			
1365	1370		1375
His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu			
1380	1385		1390

SUBSTITUTE SHEET (RULE 26)

- 140 -

Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp  
 1395 1400 1405  
 Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe  
 1410 1415 1420  
 Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr  
 1425 1430 1435 1440  
 Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met  
 1445 1450 1455  
 Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile  
 1460 1465 1470  
 Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly  
 1475 1480 1485  
 Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro  
 1490 1495 1500  
 Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg  
 1505 1510 1515 1520  
 Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro  
 1525 1530 1535  
 Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His  
 1540 1545 1550  
 Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met  
 1555 1560 1565  
 Ile Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu  
 1570 1575 1580  
 Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val  
 1585 1590 1595 1600  
 Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala  
 1605 1610 1615  
 Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg  
 1620 1625 1630  
 Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala  
 1635 1640 1645  
 Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val  
 1650 1655 1660

SUBSTITUTE SHEET (RULE 26)

- 141 -

Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
 1665 1670 1675 1680  
 Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
 1685 1690 1695  
 Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
 1700 1705 1710  
 Met Ser Ile Lys Asp Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
 1715 1720 1725  
 Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
 1730 1735 1740  
 Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn  
 1745 1750 1755 1760  
 Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
 1765 1770 1775  
 Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
 1780 1785 1790  
 Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
 1795 1800 1805  
 Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
 1810 1815 1820  
 Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
 1825 1830 1835 1840  
 Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp  
 1845 1850 1855  
 Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr  
 1860 1865 1870  
 Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys  
 1875 1880 1885  
 Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
 1890 1895 1900  
 Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
 1905 1910 1915 1920  
 Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser Tyr  
 1925 1930 1935  
 Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser

SUBSTITUTE SHEET (RULE 26)

- 142 -

1940	1945	1950
Thr Glu Ser Tyr Leu Val Met 1955	Thr Asp Leu Lys Ala Asn Arg Leu Met 1960	
Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr 1970		1980
Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys 1985	1990	1995 2000
Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Gly Ile Asn Pro 2005	2010	2015
Ile Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly 2020	2025	2030
Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp 2035	2040	2045
Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr 2050	2055	2060
Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met 2065	2070	2075 2080
Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile 2085	2090	2095
Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly 2100	2105	2110
Asn Arg Lys Leu Ile Asn Arg Phe Ile Gln Asn Leu Lys Ser Gly Tyr 2115	2120	2125
Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys 2130	2135	2140
Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val 2145	2150	2155 2160
Phe Lys Val Thr Ile Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly 2165	2170	2175
Tyr Ser Ala Leu Ile Lys Asp 2180		

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs  
(B) TYPE: nucleic acid

SUBSTITUTE SHEET (RULE 26)

- 143 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ACCAAACAAA GTTGGGTAAG GATAGATCAA TCAATGATCA TATTCTAGTA CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TGAGGAGCCT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATTCCTGGA GATTCCTCAA	240
TTACCACTCG ATCCAGACTA CTGGACCGGT TGGTCAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATCAGGCTG TTAGAGGTTG	420
TTCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAAGCAG TAGTGATCAA TCCAGGTCCG	540
GATGGTTCGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGATCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATTCTA GCCCAGATCT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC TTTACGCCGA TTCATGGTGG CTCTAATCCT GGATATCAAG AGGACACCCG	840
GGAACAAACC TAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCTTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GA CTG CATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAT CTTTACCAGC	1020
AAATGGGAGA AACTGCACCC TACATGGTAA TCCTAGAGAA CTCAATTCAG AACAAGTTCA	1080
GCGCAGGATC ATACCCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGTCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTGAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCCGAAC TCG	1260

SUBSTITUTE SHEET (RULE 26)

- 144 -

GTATCACTGC CGAGGATGCA AGGCTTGTTC CAGAGATTGC AATGCATACT ACTGAGGACA	1320
GGATCAGTAG AGCGGTCGGA CCCAGACAAG CCCAAGTATC ATTTCTACAC GGTGATCAAA	1380
GTGAGAATGA GCTACCAGGA TTGGGGGGCA AGGAAGACAG GAGGGTCAAA CAGAGTCGGG	1440
GAGAAGCCAG GGAGAGCTAC AGAGAAACCG AGTCCAGCAG AGCAAGTGAT GCGAGAGCTG	1500
CCCATCCTCC AACCAGCATG CCCCTAGACA TTGACACTGC ATCGGAGTCA GGCCAAGATC	1560
CGCAGGACAG TCGAAGGTCA GCTGACGCTC TGCTCAGGCT GCAAGCCATG GCAGGAATCT	1620
TGGAAGAACA AGGCTCAGAC ACGGACACCC CTAGGGTATA CAATGACAGA GATCTTCTAG	1680
ATTAGGTGCG AGAGGCCGAG GACCAGAACA ACATCCGCTT ACCCTCCATC ATTGTTATAA	1740
AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCAACCA ACCATCCACT CCCACGACTG	1800
GAGCCGATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGCT	1860
CTCAAGGCCG AGCCCATCGG CTCCTGGCC GTCGAGGAAG CCATGGCAGC ATGGTCAGAA	1920
ATATCAGACA ATCCAGGACA GGACCGAGCC GCCTGCAAGG AAGAGGAGGC AGGCAGTTCG	1980
GGTCTCAGCA AACCATGCTT CTCAGCAATT GGATCAACTG AAGGCGGTGC ACCTCGCATC	2040
CGCGGTGAGG GATCTGGAGA AAGCGATGAC GACGCTGAAA CTTTGGGAAT CCCCTCAAGA	2100
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTTATCATG TTTATGATCA CAGCGGTGAA	2160
GCGGTAAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT	2220
AGCACCTCT CAGGAGGAGA CGATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT	2280
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTGAGG	2340
GCTTCTGATG TTGAAACTGC AGAAGGAGGG GAGATCCAGG AGCTCCTGAA ACTCCAATCC	2400
AGAGGCAACA ACTTTCCGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC CCCGAACCCC	2460
AGTAGGGCCA GCACTTCCGA GACACCCATT AAAAAGGGGA CAGACGCGAG ATTGGCCTCA	2520
TTTGGAAACG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA	2580
CCCTCGGAAC CGTCAGGGCC AGATGCACCT GCGGGGAATG TCCCCGAGTG TGTGAGCAAT	2640
GCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCCAG	2700
AATAATGAAG AAGGGGGAGA CTATTATGAT GATGAGCTGT TCTCCGATGT CCAAGACATC	2760
AAAACAGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCTCCAA GCTAGAATCA	2820

SUBSTITUTE SHEET (RULE 26)



- 145 -

TTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAGAAGC AGATCAACAG GCAAAATATC	2880
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TTGCCATTCC TGGACTTGGG	2940
AAGGATCCCA ACGACCCAC TGCAGATGTC GAACTCAATC CCGACCTGAA ACCCATCATA	3000
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AGCCCGTTGC CAGCCGACAA	3060
CTCCAGGGAA TGAATAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTCAA	3120
CTAAAGCCGA TCGGGAAAAA GGTGAGCTCA GCCGTCGGGT TTGTCCCTGA CACCGGCCCT	3180
GCATCACGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG	3240
CGTTACCTGT TGACTCTCCT TGATGATATC AAAGGAGCCA ACGATCTTGC CAAGTTCCAC	3300
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG	3360
CCAGTCGACC TAATTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT	3420
GCCTCCTAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA	3480
AAGGGTCGAT CGCTCCGATA CAACCTACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG	3540
TCAGAGTCAT AGATCCTGGT CTAGGTGATA GGAAGGATGA ATGCTTTATG TACATGTTTC	3600
TGCTGGGGGT TGTGAGGAC AGAGATCCCC TAGGGCCTCC AATCGGGCGA GCATTGCGGT	3660
CCCTGCCCTT AGGTGTTGGT AGATCCACAG CAAAACCCGA GGAAGTCTC AAAGAGGCCA	3720
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA	3780
ACAACACCCC ACTAACCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT	3840
TCAATGCAAA CCAAGTGTGC AATGCGGTTA ATCTAATACC GCTGGACACC CCGCAGAGGT	3900
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTCCCA	3960
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTA GTGACCCTCA	4020
GGATTGACAA GGCATTGGC CCTGGGAAGA TCATCGACAA TGCAGAGCAA CTTCTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCATTGGT GGGATAGGGG	4200
GCACCACTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACTCG	4260
GGTTCAAGAA GACCTTATGT TACCCACTGA TGGATATCAA TGAAGACCTT AATCGGTTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC	4380

SUBSTITUTE SHEET (RULE 26)

- 146 -

AAGAATTCCG CATTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC 4440  
TGTAGACCGT AGTGCCCGAGC AATACCCGAA AACGACCCCC CTCATAATGA CAGCCAGAAG 4500  
GCCCCGACAA AAAAGCCCCC TCCAAAAGAC TTCACGGACC AAGCGAGAGG CCAGCCAGCA 4560  
GCCGACAGCA AGTGTGGACA CCAGGCGGCC CAAGCACAGA ACAGCCCCGA CACAAGGCCA 4620  
CCACCAGCCA TCCCAATCCG CGTCCTCCTC GTAGGACCCC CGAGGACCAA CCCCCAAGGT 4680  
CGCTCCGGAC ACAGACCACC AGCCGCATCC CCACAGCCCT CGGGAAGGA ACCCCAGCA 4740  
ACTGGAAGGC CCTTCCCCC CTCCCCAAC GCAAGAACC CACAACCGAA CCGCACAAGC 4800  
GACCGAGGTG ACCCAACCGC AGGCATCCGA CTCCTAGAC AGACCCTCC TCCCCGGCAT 4860  
ACTAAACAAA ACTTAGGGCC AAGGAACACA CACACCCGAC AGAACCAGCA CCCCAGCCCC 4920  
CGGCACCGCG CCCCCACCCC CCGAAAACCA GAGGGAGCCC CCAACCAATC CCGCCGCCCC 4980  
CCCCGGTGCC CACAGGTAGG CACACCAACC CCCGAACAGA CCCAGCACCC AGCCACCGAC 5040  
AATCCAAGAC GGGGGGCCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCATCGC 5100  
GAGGAAGCCC ACCCACCCTA CACACGACCA CGGCAACCAA ACCAGAGCCC AGACCACCTT 5160  
GGGCCACCAG CTCCAGACT CGGCCATCAC CCCGAAAAAA GGAAAGGCCA CAACCCGCGC 5220  
ACCCAGGCC CGATCCGGCG GGAAGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC 5280  
TGGGGGACCC CCAAACCGCA AAAGACATCA GTATCCCACC GCCTCTCCAA GTCCCCGGT 5340  
CTCCTCCTCT TCTCGAAGGG ACCAAAAGAT CAATCCACCA CATCCGACGA CACTCAATTC 5400  
CCCACCCCTA AAGGAGACAC CGGGAATCCC AGAATCAAGA CTCATCCAAT GTCCATCATG 5460  
GGTCTCAAGG TGAATGTCTT TGCCATATTC ATGGCAGTAC TGTTAACTCT CCAAACACCC 5520  
ACCGGTCAAA TCCATTGGGG CAATCTCTCT AAGATAGGGG TGGTAGGGAT AGGAAGTGCA 5580  
AGCTACAAAG TTATGACTCG TTCCAGCCAT CAATCATTGG TCATAAAATT AATGCCCAAT 5640  
ATAACTCTCC TCAATAACTG CACGAGGGTA GAAATTGCAG AATACAGGAG ACTACTGAGA 5700  
ACAGTTTGG AACCAATTAG AGATGCACTT AATGCAATGA CCCAGAATAT AAGACCGGTT 5760  
CAGAGTGTAG CTTCAAGTAG GAGACACAAG AGATTGCGG GAGTTGTCTT GGCAGGTGCG 5820  
GCCCTAGGCG TTGCCACAGC TGCTCAGATA ACAGCCGGCA TTGCACTTCA CCAGTCCATG 5880  
CTGAACTCTC AAGCCATCGA CAATCTGAGA GCAAGCCTGG AACTACTAA TCAGGCAATT 5940

SUBSTITUTE SHEET (RULE 26)

- 147 -

GAGGCAATCA GGCAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC	6000
ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTAAAT CGGCCAGAAG	6060
CTAGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CATTATTTGG CCCCAGCTTA	6120
CGGGACCCCA TATCTGCGGA GATATCCATC CAGGCTTTGA GCTATGCGCT TGGGGGAGAT	6180
ATCAATAAGG TATTAGAAAA GCTCGGATAC AGTGGAGGTG ATTTACTGGG CATCTTAGAG	6240
AGCAGAGGAA TAAAGGCCCG GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTCCTC	6300
AGTATAGCCT ATCCGACGCT GTCCGAGATT AAGGGGGTGA TTGTCCACCG GCTAGAGGGG	6360
GTCTCGTACA ATATAGGCTC TCAAGAGTGG TATACCACTG TGCCCAAGTA TGTTGCAACC	6420
CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT	6480
GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TCCACCAAGT CCTGTGCTCG TACACTCGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTCT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCC	6720
GTGGTCGAGG TGAACGGTGT GACCATCCAA GTCGGGAGCA GGAGGTATCC GGACGCGGTG	6780
TACCTGCACA GAATTGACCT CGGTCTCCC ATATCATTGG AGAAGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGCTGGAG GATGCCAAGG AATTGCTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTAGCA TAGTTTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACAGGG	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCCCTACAA CTCTTGAAAC ACAGATTTCC	7140
CACAAGTCTC CTCTCCGTCA TCAAGCAACC ACCGCATCCA GCATCAAGGC CACCCGAAAT	7200
TGTCTCCGGC TTCCCTCTGG CCGAACGATA TCGGTAGTTA ATTAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC ACCGAGACCG AATAAATGCC TTCTACAAAG ACAACCCCCA	7320
TCCTAAGGGA AGTAGGATAG TTATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTATTCG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTC CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAGAGCC TCAGCACCAA	7500

SUBSTITUTE SHEET (RULE 26)

- 148 -

TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
GATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTCACTGACC TAGTGAAATT	7620
CATCTCTGAC AAAATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAACTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGG CCAGGGCAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCCTGTTGGA CTTGTATTTA AGTCGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACCATGACAT CCCAGGGAAT GTACGGGGGA ACTTACCTAG TGGGAAAGCC	7980
TAATCTGAGC AGTAAAGGGT CAGAGTTGTC ACAACTGAGC ATGCACCGAG TGTTTGAAGT	8040
AGGGGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATTTTGA	8100
GCAACCAGTC AGTAATGATT TCAGCAACTG CATGGTGGCT TTGGGGGAGC TCAGGTTCCG	8160
AGCCCTCTGT CACAGGGAAG ATTCTGTCAC GGTTCCTAT CAGGGGTCAG GGAAAGGTGT	8220
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CCCCCTATCA ACGGATGATC CAGTGATAGA TAGGCTTTAC CTCTCATCTC ACAGAGGTGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGGACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAGC AGGCGTGTA GGGTAAAAAC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTAATCT	8520
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCAGGA TTCGGGCCAT TGATCACACA	8580
CGGTTTCAGG ATGGACCTAT ACAAACCAA CCACAACAAT GTGTATTGGC TGACTATCCC	8640
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTTAGTCCC AACCTCTTCA CTGTTCCAAT CAAGGAAGCA GCGGAGGACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAACCT AGTTCCAATC TGGTAATTCT	8820
ACCTGGTCAG GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCCAGGG TTGAACATGC	8880
TGTGGTTTAT TATGTTTACA GCCCAGGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT	8940
GCCTATAAAG GGGGTCCCAA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA TTCAGAATCT GGTGGACATA TCACTCACTC	9060

SUBSTITUTE SHEET (RULE 26)

- 149 -

TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACTCGG GAAGATGGAA CCAATCGCAG	9120
ATAGGGCTGC CAGTGAACCG ATCACATGAT GTCCTCAGA CACCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTTCCC GTCATGGACT	9240
CGCTATCTGT CAACCAGATC TTGTACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCTATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTTGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACC GCC TAAAAAACGG ATTCTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCACT	9480
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540
CAAGGAAGAT CCGTGAGCTC CTAAAAAAGG GAAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCCTGAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAATT ATTAACCTGG GAGTTTACAT GCACAGCTCC CAATGGTTTG	9720
AGCCCTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGC	9840
TGTTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAGGA GTCTCAACAT GTATATTACC	9900
TGACGTTTGA ACTGGTTTGG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC CATTGATGCT AGGTATGCAG AACTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCTATGC	10080
TGGAGCCACT TTCCTTGCT TACCTGCAAC TGAGGGACAT AACAGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCATGAG TTAATTGAAG CCTTAGATTA CATTTCATA ACTGATGACA	10260
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTCAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAGGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCCCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAGATCAT	10560
TTGCTGGAGT GAGATTTGGC TGTTTTATGC CTCTTAGCCT GGACAGTGAT CTGACAATGT	10620

SUBSTITUTE SHEET (RULE 26)

- 150 -

ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG	10680
AGTTCCTGCG TTACGATCCT CCCAAGGGAA CCGGGTCACG GAGGCTTGTA GATGTTTTCC	10740
TTAATGATTC GAGCTTTGAC CCATATGATA TGATAATGTA TGTCGTAAGT GGAGCCTACC	10800
TCCATGACCC TGAGTTCAAT CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG	10860
GTAGACTTTT CGCTAAAATG ACTTACAAAA TGAGGGCATG CCAAGTGATC GCTGAAAATC	10920
TAATCTCAA CGGGATTGGC AAGTATTTTA AGGACAATGG GATGGCCAAG GATGAGCAGG	10980
ATTTGACTAA GGCACTCCAC ACTCTGGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA	11040
GTCACAGGGG GGGGCCAGTC TTA AAAACCT ACTCCCGAAG CCCAGTCCAC ACAAGTACCA	11100
GGAACGTAA AGCAGAAAAA GGGTTTGTAG GATTCCCTCA TGTAATTCGG CAGAATCAAG	11160
ACACTGATCA TCCGGAGAAT ATAGAAACCT ACGAGACAGT CAGCGCATTT ATCAGCACTG	11220
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTATTT GCACAGAGGC	11280
TAAATGAGAT TTACGGATTA CCCTCATTTT TTCAGTGGCT GCATAAGAGG CTTGAAACCT	11340
CTGTCCTCTA TGTAAGTGAT CCTCATTGCC CCCCCGACCT TGACGCCCAT GTCCCGTTAT	11400
GCAAAGTCCC CAATGACCAA ATCTTCATCA AGTACCCTAT GGGAGGTATA GAAGGGTATT	11460
GTCAGAAGCT GTGGACCATC AGCACCATTG CCTACTTATA CCTGGCTGCT TATGAGAGCG	11520
GGGTAAGGAT TGCCTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG	11580
TACCCAGCAC ATGGCCTTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT	11640
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ACATTGGCCA TCACCTCAAG GCAAATGAGA	11700
CAATTGTTTT ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG	11760
TGTCCCAATC ACTCAAGAGC ATTGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG	11820
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT	11880
ATGACCGTTA TCTTGATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTTTGATCT	11940
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCGAGATGT AGTCATACCC CTCCTCACAA	12000
ACAACGATCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC	12060
TGAACATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG	12120
ATCTCAAGAG AATGATTCTC GCATCACTAA TGCCTGAAGA GACCCTCCAT CAAGTAATGA	12180

SUBSTITUTE SHEET (RULE 26)

- 151 -

CACAACAACC	GGGGGACTCT	TCATTCTAG	ACTGGGCTAG	CGACCCTTAC	TCAGCAAATC	12240
TTGTATGCGT	CCAGAGCATC	ACTAGACTCC	TCAAGAACAT	AACTGCAAGG	TTTGTCTTAA	12300
TCCATAGTCC	AAACCCAATG	TTAAAAGGGT	TATTCCATGA	TGACAGTAA	GAAGAGGACG	12360
AGAGACTGGC	GGCATTCTCT	ATGGACAGGC	ATATTATAGT	ACCTAGGGCA	GCTCATGAAA	12420
TCCTGGATCA	TAGTGTCA	GGGGCAAGAG	AGTCTATTGC	AGGCATGCTA	GATACCACAA	12480
AAGGCCTGAT	TCGAGCCAGC	ATGAGGAAGG	GGGGGTAAAC	CTCTCGAGTG	ATAACCAGAT	12540
TGTCCAATTA	TGACTATGAA	CAATTTAGAG	CAGGGATGGT	GCTATTGACA	GGAAGAAAGA	12600
GAAATGTCTT	CATTGACAAA	GAGTCATGTT	CAGTGCAGCT	GGCTAGAGCC	CTAAGAAGCC	12660
ATATGTGGGC	AAGACTAGCT	CGAGGACGGC	CTATTTACGG	CCTTGAGGTC	CCTGATGTAC	12720
TAGAATCTAT	GCGAGGCCAC	CTTATTCGGC	GTCATGAGAC	ATGTGTCATC	TGCGAGTGTG	12780
GATCAGTCAA	CTACGGATGG	TTTTTTGTCC	CCTCGGGTTG	CCAACTGGAT	GATATTGACA	12840
AGGAAACATC	ATCCTTGAGA	GTCCCATATA	TTGGTTCTAC	CACTGATGAG	AGAACAGACA	12900
TGAAGCTTGC	CTTCGTAAGA	GCCCCAAGTA	GATCCTTGCG	ATCTGCCGTT	AGAATAGCAA	12960
CAGTGTACTC	ATGGGCTTAC	GGTGATGATG	ATAGCTCTTG	GAACGAAGCC	TGGTTGTTGG	13020
CAAGGCAAAG	GGCCAATGTG	AGCCTGGAGG	AGCTAAGGGT	GATCACTCCC	ATCTCGACTT	13080
CGACTAATTT	AGCGCATAGG	TTGAGGGATC	GTAGCACTCA	AGTGAAATAC	TCAGGTACAT	13140
CCCTTGTCAG	AGTGGCAAGG	TATACCACAA	TCTCCAACGA	CAATCTCTCA	TTTGTCTAT	13200
CAGATAAGAA	AGTTGATACT	AACTTTATAT	ACCAACAAGG	AATGCTTCTA	GGGTTGGGTG	13260
TTTTAGAAAC	ATTGTTTCGA	CTCGAGAAAG	ATACTGGATC	ATCTAACACG	GTATTACATC	13320
TTCACGTCGA	AACAGATTGT	TGCGTGATCC	CGATGATAGA	TCATCCCAGG	ATACCCAGCT	13380
CCC GCAAGCT	AGAGCTGAGG	GCAGAGCTAT	GTACCAACCC	ATTGATATAT	GATAATGCAC	13440
CTTTAATTGA	CAGAGATGCA	ACAAGGCTAT	ACACCCAGAG	CCATAGGAGG	CACCTTGTGG	13500
AATTTGTTAC	ATGGTCCACA	CCCCAACTAT	ATCACATTCT	AGCTAAGTCC	ACAGCACTAT	13560
CTATGATTGA	CCTGGTAACA	AAATTTGAGA	AGGACCATAT	GAATGAAATT	TCAGCTCTCA	13620
TAGGGGATGA	CGATATCAAT	AGTTTCATAA	CTGAGTTTCT	GCTTATAGAG	CCAAGATTAT	13680
TCACCATCTA	CTTGGGCCAG	TGTGCAGCCA	TCAATTGGGC	ATTTGATGTA	CATTATCATA	13740

SUBSTITUTE SHEET (RULE 26)

- 152 -

GACCATCAGG GAAATATCAG ATGGGTGAGC TGTGTCTTC GTTCCTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGACACA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTTGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTTTTGT GTGAAAGCGA TGAGGATGTA GTACCGGACA 14040  
GATTCGACAA CATCCAGGCA AAACACTTGT GTGTTCTGGC AGATTTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCGATTGCA GGTCTAAGGC CGGTAGAGAA ATGTGCAGTT CTAACCGATC 14160  
ATATCAAGGC AGAGGCTAGG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTAATCATGC TCTCTGACTT ATCTCCGTCG AGGATCTATC AAACAGATAA 14280  
GATTGAGAGT TGATCCAGGA TTCATTTTTG ATGCCCTCGC TGAGGTAAAT GTCAGTCAGC 14340  
CAAAGGTCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGATTTCAGA CCTCCACACG 14400  
ATGATGTTGC AAAATTGCTC AAAGATATCA ACACAAGCAA GCACAATCTT CCCATTTTCAG 14460  
GGGGTAGTCT TGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTA AACTCATCTG 14520  
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAAG 14580  
ACGGCTTGTT CTTGGGTGAG GGGTCGGGTT CTATGTTGAT CACTTATAAG GAGATACTAA 14640  
AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG 14700  
AATTAGCACC CTATCCCTCC GAAGTTGGCC TTGTGCAACA CAGAATGGGA GTAGGTAATA 14760  
TTGTCAAGGT GCTCTTTAAC GGGAGGCCCG AAGTCACGTG GGTAGGCAGT ATAGATTGCT 14820  
TCAATTTTCAT AGTCAGTAAT ATCCCTACCT CTAGTGTGGG ATTTATCCAT TCAGATATAG 14880  
AGACCTTACC CAACAAAGAT ACTATAGAGA AGTTAGAGGA ATTGGCAGCC ATCTTATCGA 14940  
TGGCTCTACT CCTTGGCAAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTCAGCG 15000  
GGGATTTTGT TCAGGGATTT ATAAGCTATG TAGGGTCTCA TTATAGAGAA GTGAACCTTG 15060  
TCTACCCTAG GTACAGCAAC TTCATATCTA CTGAATCTTA TTTAGTTATG ACAGATCTCA 15120  
AAGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGC 15180  
GGACTTCACC TGGACTTATA GGTCACATCC TATCTATCAA GCAACTAAGC TGCATACAAG 15240  
CAATTGTGGG AGGCGCAGTT AGTAGAGGTG ATATCAACCC TATTCTGAAA AACTTACAC 15300

SUBSTITUTE SHEET (RULE 26)



CTATAGAGCA	GGTGCTGATC	AGTTGCGGGT	TGGCAATTAA	CGGACCCTAAG	CTGTGCAAAAG	15360
AATTAATCCA	CCATGATGTT	GCCTCAGGGC	AAGATGGATT	GCTTAACTCT	ATACTCATCC	15420
TCTACAGGGA	GTTGGCAAGA	TTCAAAGACA	ACCAAAGAAG	TCAACAAGGG	ATGTTCCACG	15480
CTTACCCCGT	ATTGGTAAGT	AGTAGGCAAC	GAGAACTTGT	ATCTAGGATC	ACTCGCAAAT	15540
TTTGGGGGCA	TATTCTTCTT	TACTCCGGGA	ACAGAAAGTT	GATAAATCGG	TTTATCCAGA	15600
ATCTCAAGTC	CGGTTATCTA	ATACTAGACT	TACACCAGAA	TATCTTCGTT	AAGAATCTAT	15660
CCAAGTCAGA	GAAACAGATT	ATTATGACGG	GGGGTTTAAA	ACGTGAGTGG	GTTTTTAAGG	15720
TAACAGTCAA	GGAGACCAAA	GAATGGTATA	AGTTAGTCGG	ATACAGCGCT	CTGATTAAGG	15780
ATTAATTGGT	TGAACTCCGG	AACCCTAATC	CTACCCTAGG	TAGTTAGGCA	TTATTTGCAA	15840
TATATTAAAG	AAAACCTTGA	AAATACGAAG	TTTCTATTCC	CAGCTTTGTC	TGGT	15894

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
1 5 10 15

Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
20 25 30

Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
35 40 45

Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
50 55 60

Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
65 70 75 80

Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn  
85 90 95

**SUBSTITUTE SHEET (RULE 26)**

- 154 -

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys  
 100 105 110  
 Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu  
 115 120 125  
 Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp  
 130 135 140  
 Ile Lys Glu Lys Ile Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln  
 145 150 155 160  
 Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg  
 165 170 175  
 Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr  
 180 185 190  
 Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
 195 200 205  
 Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
 210 215 220  
 Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
 225 230 235 240  
 Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Ala Glu Leu Leu Gly  
 245 250 255  
 Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
 260 265 270  
 Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu  
 275 280 285  
 Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe  
 290 295 300  
 Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly  
 305 310 315 320  
 Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr  
 325 330 335  
 Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe  
 340 345 350  
 Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu  
 355 360 365

SUBSTITUTE SHEET (RULE 26)

- 155 -

Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr  
 370 375 380  
 Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr  
 385 390 395 400  
 Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His  
 405 410 415  
 Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr  
 420 425 430  
 His Glu Gln Cys Val Asp Asn Trp Arg Ser Phe Ala Gly Val Arg Phe  
 435 440 445  
 Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu  
 450 455 460  
 Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr  
 465 470 475 480  
 Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg  
 485 490 495  
 Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp  
 500 505 510  
 Met Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe  
 515 520 525  
 Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg  
 530 535 540  
 Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala  
 545 550 555 560  
 Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly  
 565 570 575  
 Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala  
 580 585 590  
 Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro  
 595 600 605  
 Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn  
 610 615 620  
 Val Lys Ala Glu Lys Gly Phe Val Gly Phe Pro His Val Ile Arg Gln  
 625 630 635 640  
 Asn Gln Asp Thr Asp His Pro Glu Asn Ile Glu Thr Tyr Glu Thr Val

SUBSTITUTE SHEET (RULE 26)

- 156 -

645	650	655
Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg		
660	665	670
Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly		
675	680	685
Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val		
690	695	700
Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Val		
705	710	715
Pro Leu Cys Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met		
725	730	735
Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile		
740	745	750
Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser		
755	760	765
Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro		
770	775	780
Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr		
785	790	795
Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His		
805	810	815
His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr		
820	825	830
Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys		
835	840	845
Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr		
850	855	860
Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu		
865	870	875
Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val		
885	890	895
Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met		
900	905	910
Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile		
915	920	925

SUBSTITUTE SHEET (RULE 26)

- 157 -

Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940

Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960

Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu  
 965 970 975

Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990

Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005

Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His  
 1010 1015 1020

Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
 1025 1030 1035 1040

Glu Asp Glu Arg Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
 1045 1050 1055

Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
 1060 1065 1070

Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
 1075 1080 1085

Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
 1090 1095 1100

Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly  
 1105 1110 1115 1120

Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu  
 1125 1130 1135

Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg  
 1140 1145 1150

Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly  
 1155 1160 1165

His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser  
 1170 1175 1180

Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp  
 1185 1190 1195 1200

SUBSTITUTE SHEET (RULE 26)

- 158 -

Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr  
 1205 1210 1215  
 Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser  
 1220 1225 1230  
 Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala  
 1235 1240 1245  
 Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg  
 1250 1255 1260  
 Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile  
 1265 1270 1275 1280  
 Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln  
 1285 1290 1295  
 Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr  
 1300 1305 1310  
 Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp  
 1315 1320 1325  
 Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu  
 1330 1335 1340  
 Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val  
 1345 1350 1355 1360  
 Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp  
 1365 1370 1375  
 His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu  
 1380 1385 1390  
 Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp  
 1395 1400 1405  
 Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe  
 1410 1415 1420  
 Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr  
 1425 1430 1435 1440  
 Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met  
 1445 1450 1455  
 Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile  
 1460 1465 1470  
 Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly

SUBSTITUTE SHEET (RULE 26)

- 159 -

1475	1480	1485
Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro 1490	1495	1500
Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg 1505	1510	1515 1520
Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro 1525	1530	1535
Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His 1540	1545	1550
Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met 1555	1560	1565
Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu 1570	1575	1580
Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val 1585	1590	1595 1600
Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala 1605	1610	1615
Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg 1620	1625	1630
Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala 1635	1640	1645
Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val 1650	1655	1660
Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys 1665	1670	1675 1680
Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala 1685	1690	1695
Glu Val Asn Val Ser Gln Pro Lys Val Gly Ser Asn Asn Ile Ser Asn 1700	1705	1710
Met Ser Ile Lys Asp Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu 1715	1720	1725
Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly 1730	1735	1740
Ser Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn 1745	1750	1755 1760

SUBSTITUTE SHEET (RULE 26)

- 160 -

Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
 1765 1770 1775  
 Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
 1780 1785 1790  
 Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
 1795 1800 1805  
 Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
 1810 1815 1820  
 Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
 1825 1830 1835 1840  
 Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp  
 1845 1850 1855  
 Val Gly Ser Ile Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr  
 1860 1865 1870  
 Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys  
 1875 1880 1885  
 Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
 1890 1895 1900  
 Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
 1905 1910 1915 1920  
 Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His  
 1925 1930 1935  
 Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser  
 1940 1945 1950  
 Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met  
 1955 1960 1965  
 Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr  
 1970 1975 1980  
 Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys  
 1985 1990 1995 2000  
 Ile Gln Ala Ile Val Gly Gly Ala Val Ser Arg Gly Asp Ile Asn Pro  
 2005 2010 2015  
 Ile Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Ser Cys Gly  
 2020 2025 2030

SUBSTITUTE SHEET (RULE 26)



- 161 -

Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp  
 2035 2040 2045  
 Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr  
 2050 2055 2060  
 Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met  
 2065 2070 2075 2080  
 Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Val  
 2085 2090 2095  
 Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly  
 2100 2105 2110  
 Asn Arg Lys Leu Ile Asn Arg Phe Ile Gln Asn Leu Lys Ser Gly Tyr  
 2115 2120 2125  
 Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
 2130 2135 2140  
 Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
 2145 2150 2155 2160  
 Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
 2165 2170 2175  
 Tyr Ser Ala Leu Ile Lys Asp  
 2180

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACCAAACAAA GTTGGGTAAG GATAGTTCAA TCAATGATCA TCTTCTAGTG CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCCTCAA	240

- 162 -

TTACCACTCG ATCCAGACTT CTGGACCGGT TGGTCAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAATTAG TAGTGATCAA TCCAGGTTG	540
GATGGTTCGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCCCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TGGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGAACACCCG	840
GAAACAAACC CAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GA CTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAC CTTTACCAGC	1020
AAATGGGGGA AACTGCACCC TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAAGTTCA	1080
GTGCAGGATC ATACCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGCCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTAAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCTGAACTCG	1260
GTATCACTGC CGAGGATGCA AGGCTTGTTT CAGAGATTGC AATGCATACT ACTGAGGACA	1320
AGATCAGTAG AGCGGTTGGA CCCAGACAAG CCCAAGTATC ATTTCTACAC GGTGATCAAA	1380
GTGAGAATGA GCTACCGAGA TTGGGGGGCA AGGAAGATAG GAGGGTCAAA CAGAGTCGAG	1440
GAGAAGCCAG GGAGAGCTAC AGAGAAACCG GGCCAGCAG AGCAAGTGAT GCGAGAGCTG	1500
CCCATCTTCC AACCGGCACA CCCCTAGACA TTGACACTGC AACGGAGTCC AGCCAAGATC	1560
CGCAGGACAG TCGAAGGTCA GCTGACGCCC TGCTTAGGCT GCAAGCCATG GCAGGAATCT	1620
CGGAAGAACA AGGCTCAGAC ACGGACACCC CTATAGTGTA CAATGACAGA AATCTTCTAG	1680
ACTAGGTGCG AGAGGCCGAG GGCCAGAACA ACATCCGCCT ACCCTCCATC ATTGTTATAA	1740
AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCCATCA ACCATCCACT CCCACGATTG	1800

SUBSTITUTE SHEET (RULE 26)

- 163 -

GAGCCAATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGCT 1860  
CTCAAGGCCG AGCCCATCGG CTCCTGGCC ATCGAGGAAG CTATGGCAGC ATGGTCAGAA 1920  
ATATCAGACA ACCCAGGACA GGAGCGAGCC ACCTGCAGGG AAGAGAAGGC AGGCAGTTCC 1980  
GGTCTCAGCA AACCATGCCT CTCAGCAATT GGATCAACTG AAGGCGGTGC ACCTCGCATC 2040  
CGCGGTGAGG GACCTGGAGA GAGCGATGAC GACGCTGAAA CTTTGGGAAT CCCCCAAGA 2100  
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTATTACG TTTATGATCA CAGCGGTGAA 2160  
GCGGTTAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT 2220  
AGCACCTCT CAGGAGGAGA CAATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT 2280  
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTGAGG 2340  
GCTTCTGATG TTGAACTGC AGAAGGAGGG GAGATCCAG AGCTCCTGAG ACTCCAATCC 2400  
AGAGGCAACA ACTTTCCGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC CCCGGACCCC 2460  
GGTAGGGCCA GCACTTCCGG GACACCCATT AAAAAGGGCA CAGACGCGAG ATTAGCCTCA 2520  
TTTGGAACGG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCAATGTGC TCGAAAGTCA 2580  
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GCGGGGAATG TCCCCGAGTG TGTGAGCAAT 2640  
GCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCCAG 2700  
AATAATGAAG AAGGGGGAGA CTATTATGAT GATGAGCTGT TCTCTGATGT CCAAGATATT 2760  
AAAACAGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCTCCA GCTAGAATCA 2820  
CTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAGAAGC AGATCAACAG GCAAAATATC 2880  
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG 2940  
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA 3000  
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCTTGC CAGCCGACAA 3060  
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTGAG 3120  
CTAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCCCTGA CACCGGCCCT 3180  
GCATCAGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG 3240  
CGTTACCTGA TGAATCTCCT TGATGATATC AAAGGAGCCA ATGATCTTGC CAAGTTCCAC 3300  
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG 3360

SUBSTITUTE SHEET (RULE 26)

- 164 -

CCAGTCGACC CAACTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT	3420
GCCTCCCAAG TTCCACAATG ACAGAGACCT ACGACTTCGA CAAGTCGGCA TGGGACATCA	3480
AAGGGTCGAT CGCTCCGATA CAACCCACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG	3540
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTATG TACATGTTTC	3600
TGCTGGGGGT TGTTGAGGAC AGCGATTCCC TAGGGCCTCC AATCGGGCGA GCATTGGGT	3660
CCCTGCCCTT AGGTGTTGGC AGATCCACAG CAAAGCCCCG AAAACTCCTC AAAGAGGCCA	3720
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA	3780
ACAACACCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT	3840
TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTCGATACC CCGCAGAGGT	3900
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGATAA CGGGTATTAC ACCGTTCCCTA	3960
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCTTA	4020
GGATTGACAA GGCATAGGC CCTGGGAAGA TCATCGACAA TACAGAGCAA CTTCTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACCTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCACTTGGT GGGATAGGGG	4200
GCACCAGTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACCTG	4260
GGTTCAAGAA GACCTTATGT TACCCGCTGA TGGATATCAA TGAAGACCTT AATCGATTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC	4380
AAGAATTCCG CATTTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC	4440
TGTAGACCGT AGTGCCGAGC AATGCCCGAA AACGACCCCC CTCACAATGA CAGCCAGAAG	4500
GCCCGGACAA AAAAGCCCCC TCCGAAAGAC TCCACGGACC AAGCGAGAGG CCAGCCAGCA	4560
GCCGACGGCA AGCGCGAACA CCAGGCGGCC CCAGCACAGA ACAGCCCCGA CACAAGGCCA	4620
CCACCAGCCA CCCCAATCTG CATCCTCCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGC	4680
TGCCCCCGAT CCAAACCACC AACC GCATCC CCACCACCCC CGGGAAAGAA ACCCCCAGCA	4740
ATTGGAAGGC CCCTCCCCCT CTTCTCAAC ACAAGAACTC CACAACCGAA CCGCACAAGC	4800
GACCGAGGTG ACCCAACCGC AGGCATCCGA CTCCCTAGAC AGATCCTCTC TCCCCGGCAA	4860
ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAC AGAACCAGGA CCCC GGCCCA	4920

SUBSTITUTE SHEET (RULE 26)

- 165 -

CGGCGCCGCG	CCCCCAACCC	CCGACAACCA	GAGGGAGCCC	CCAACCAATC	CCGCCGGCTC	4980
CCCCGGTGCC	CACAGGCAGG	GACACCAACC	CCCGAACAGA	CCCAGCACCC	AACCATCGAC	5040
AATCCAAGAC	GGGGGGGCCC	CCCCAAAAAA	AGGCCCCCAG	GGGCCGACAG	CCAGCACCGC	5100
GAGGAAGCCC	ACCCACCCCA	CACACGACCA	CGGCAACCAA	ACCAGAACCC	AGACCACCCCT	5160
GGGCCACCAG	CTCCCAGACT	CGGCCATCAC	CCCGCAGAAA	GGAAAGGCCA	CAACCCGCGC	5220
ACCCAGCCCC	CGATCCGGCG	GGGAGCCACC	CAACCCGAAC	CAGCACCCAA	GAGCGATCCC	5280
CGAAGGACCC	CCGAACCGCA	AAGGACACCA	GTATCCCACA	GCCTCTCCAA	GTCCCCCGGT	5340
CTCCTCCTCT	TCTCGAAGGG	ACCAAAAGAT	CAATCCACCA	CACCCGACGA	CACTCAACTC	5400
CCCACCCCTA	AAGGAGACAC	CGGGAATCCC	AGAATCAAGA	CTCATCCAAT	GTCCATCATG	5460
GGTCTCAAGG	TGAACGTCTC	TGCCATATTC	ATGGCAGTAC	TGTTAACTCT	CCAAACACCC	5520
ACCGGTCAAA	TCCATTGGGG	CAATCTCTCT	AAGATAGGGG	TGGTAGGAAT	AGGAAGTGCA	5580
AGCTACAAAG	TTATGACTCG	TTCCAGCCAT	CAATCATTAG	TCATAAAATT	AATGCCCAAT	5640
ATAACTCTCC	TCAATAACTG	CACGAGGGTA	GAGATTGCAG	AATACAGGAG	ACTACTGAGA	5700
ACAGTTTTGG	AACCAATTAG	AGATGCACTT	AATGCAATGA	CCCAGAATAT	AAGACCGGTT	5760
CAGAGTGTAG	CTTCAAGTAG	GAGACACAAG	AGATTTGCGG	GAGTAGTCCT	GGCAGGTGCG	5820
GCCCTAGGCG	TTGCCACAGC	TGCTCAGATA	ACAGCCGGCA	TTGCACTTCA	CCAGTCCATG	5880
CTGAACTCTC	AAGCCATCGA	CAATCTGAGA	GCGAGCCTGG	AACTACTAA	TCAGGCAATT	5940
GAGGCAATCA	GACAAGCAGG	GCAGGAGATG	ATATTGGCTG	TTCAGGGTGT	CCAAGACTAC	6000
ATCAATAATG	AGCTGATACC	GTCTATGAAC	CAACTATCTT	GTGATTTAAT	CGGCCAGAAG	6060
CTCGGGCTCA	AATTGCTCAG	ATACTATACA	GAAATCCTGT	CATTATTTGG	CCCCAGTTTA	6120
CGGGACCCCA	TATCTGCGGA	GATATCTATC	CAGGCTTTGA	GCTATGCGCT	TGGAGGAGAC	6180
ATCAATAAGG	TGTTAGAAAA	GCTCGGATAC	AGTGGAGGTG	ATTTACTGGG	CATCTTAGAG	6240
AGCGGAGGAA	TAAAGGCCCC	GATAACTCAC	GTCGACACAG	AGTCCTACTT	CATTGTCCTC	6300
AGTATAGCCT	ATCCGACGCT	GTCCGAGATT	AAGGGGGTGA	TTGTCCACCG	GCTAGAGGGG	6360
GTCTCGTACA	ACATAGGCTC	TCAAGAGTGG	TATACCACTG	TGCCCAAGTA	TGTTGCAACC	6420
CAAGGGTACC	TTATCTCGAA	TTTTGATGAG	TCATCGTGTA	CTTTCATGCC	AGAGGGGACT	6480

SUBSTITUTE SHEET (RULE 26)

- 166 -

GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TACACCAAGT CCTGTGCTCG TACACTCGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCC	6720
GTAGTCGAGG TGAACGGCGT GATCATCCAA GTCGGGAGCA GGAGGTATCC AGACGCTGTG	6780
TACTTGACCA GAATTGACCT CGGTCCTCCC ATATCATTGG AGAGGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTAGCA TAGTCTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACGGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCACCCA GCATCAAGCC CACCTGAAAT	7200
TATCTCCGGC TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATTAAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA	7320
TCCCAAGGGA AGTAGGATAG TCATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTGTTTG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500
TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TCACTGACC TAGTGAAATT	7620
CATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCCTGTTAGA CTTGTATTTA GGTGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACTATGACAT CCCAGGGAAT GATGGGGGA ACTTACCTAG TGGAAAAGCC	7980
TAATCTGAGC AGCAAAAGGT CAGAGTTGTC ACAACTGAGC ATGTACCGAG TGTTTGAAGT	8040

SUBSTITUTE SHEET (RULE 26)

- 167 -

AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATCTTGA 8100  
GCAACCAGTC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAACCTCGC 8160  
AGCCCTTTGT CACGGGGAAG ATTCTATCAC AATTCCTTAT CAGGGATCAG GGAAAGGTGT 8220  
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT 8280  
CCCCTTATCA ACGGATGATC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT 8340  
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGAACAGATG ACAAGTTGCG 8400  
AATGGAGACA TGCTTCCAAC AGGCGTGTA GGGTAAATC CAAGCACTCT GCGAGAATCC 8460  
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTGATCT 8520  
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA 8580  
CGGTTCAGGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGAATATCCC 8640  
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA 8700  
GGTTAGTCCC TACCTCTTCA CTGTCCCAAT TAAGGAAGCA GCGGAAGACT GCCATGCCCC 8760  
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAATC AGTTCCAATC TGGTGATTCT 8820  
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCCAGGG TTGAACATGC 8880  
TGTGGTTTAT TACGTTTACA GCCCAAGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT 8940  
GCCTATAAAG GGGGTCCCCA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT 9000  
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC 9060  
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATCGCAG 9120  
ATAGGGCTGC TAGTGAACCA ATCACATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT 9180  
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT 9240  
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA 9300  
ATAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC 9360  
CTCACTGTG TCAGAACATC AAGCACCGCC TAAAAAACGG ATTTTCCAAC CAAATGATTA 9420  
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCACT 9480  
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA 9540  
CGAGGAAGAT CCGTGAACTC CTCAAAAAGG GGAATTCGCT GTACTCCAAA GTCAGTGATA 9600

SUBSTITUTE SHEET (RULE 26)

- 168 -

AGGTTTTCCA ATGCTTAAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAAGTTGG GAGTTTACAT GCACAGCTCC CAGTGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACATTGA ACTGGTTTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTCACTTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCATGAG TTAATTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA	10260
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT	10560
TTGCTGGAGT GAAATTTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT	10620
ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG	10680
AGTTCCTGCG TTACGACCCT CCCAAGGGAA CCGGGTCAGG GAGGCTTGTA GATGTTTTC	10740
TTAATGATTC GAGCTTTGAC CCATATGATG TGATAATGTA TGTTGTAAGT GGAGCTTACC	10800
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG	10860
GTAGACTTTT TGCTAAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC	10920
TAATCTCAAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCAGC	10980
ATTTGACTAA GGCACCTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA	11040
GTCACAGGGG GGGGCCAGTC TTAACCACTT ACTCCGAAG CCCAGTCCAC ACAAGTACCA	11100
GGAACGTGAG AGCAGCAAAA GGGTTTATAG GGTTCCTCA AGTAATTCGG CAGGACCAAG	11160

SUBSTITUTE SHEET (RULE 26)



- 169 -

ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATTT ATCACGACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTG CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCCTGTA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCAT ATCCCATTAT 11400  
ATAAAGTCCC CAATGATCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATTC CCTATCTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCCTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCTCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CCTTGTCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTCTGATCT 11940  
CTCTTGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCCTCACTAA TGCCTGAAGA GACCTCCAT CAAGTAATGA 12180  
CACAACAACC GGGGGACTCT TCATTCCTAG ACTGGGCTAG CGACCTTAC TCAGCAAATC 12240  
TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA 12300  
TCCATAGTCC AAACCCAATG TTAAGGAT TATCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC GGCATTCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCACA GGGGCAAGAG AGTCTATTGC AGGCATGCTG GATACCACAA 12480  
AAGGCTTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGAATATGAA CAATTCAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGC 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GCGGAGAGCT CTAAGAAGCC 12660  
ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC 12720

SUBSTITUTE SHEET (RULE 26)

- 170 -

TAGAATCTAT GCGAGGCCAC CTTATTCGGC GTCATGAGAC ATGTGTCATC TCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTGTGTC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTGTAAGA GCCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA 12960  
CAGTGTACTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG 13020  
CTAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCGCATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTCGG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT 13200  
CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAAGG AATGCTTCTA GGGTTGGGTG 13260  
TTTAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TGCCTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CCCGCAAGCT AGAGCTGAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG 13500  
AATTTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTCATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740  
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTGTGTCATC GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGACAA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTCTTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA 14040  
GATTGACAA CATCCAGGCA AAACACTTAT GTGTTCTGGC AGATTTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCAATTCGA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC 14160  
ATATCAAGGC AGAGGCTATG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTAATCATGC TCTCTGACTT ATCTCCGGCG AGGATCGATC AAACAGATAA 14280

SUBSTITUTE SHEET (RULE 26)

- 171 -

GATTGAGAGT	TGATCCAGGA	TTCATTTTCG	ACGCCCTCGC	TGAGGTAAAT	GTCAGTCAGC	14340
CAAAGATCGG	CAGCAACAAC	ATCTCAAATA	TGAGCATCAA	GGCTTTCAGA	CCCCCACACG	14400
ATGATGTTGC	AAAATTGCTC	AAAGATATCA	ACACAAGCAA	GCACAATCTT	CCCATTTCAG	14460
GGGGCAATCT	CGCCAATTAT	GAAATCCATG	CTTTCGCGAG	AATCGGGTTG	AACTCATCTG	14520
CTTGCTACAA	AGCTGTTGAG	ATATCAACAT	TAATTAGGAG	ATGCCTTGAG	CCAGGGGAGG	14580
ACGGCTTGTT	CTTGGGTGAG	GGATCGGGTT	CTATGTGAT	CACTTATAAG	GAGATACTTA	14640
AACTAAACAA	GTGCTTCTAT	AATAGTGGGG	TTTCCGCCAA	TTCTAGATCT	GGTCAAAGGG	14700
AATTAGCACC	CTATCCCTCC	GAAGTTGGCC	TTGTGGAACA	CAGAATGGGA	GTAGGTAATA	14760
TTGTCAAAGT	GCTCTTTAAC	GGGAGGCCCG	AAGTCACGTG	GGTAGGCAGT	GTAGATTGCT	14820
TCAATTTTCAT	AGTTAGTAAT	ATCCCTACCT	CTAGTGTTGG	GTTTATCCAT	TCAGATATAG	14880
AGACCTTGCC	TGACAAAGAT	ACTATAGAGA	AGCTAGAGGA	ATTGGCAGCC	ATCTTATCGA	14940
TGGCTCTGCT	CCTGGGCAAA	ATAGGATCAA	TACTGGTGAT	TAAGCTTATG	CCTTTCAGCG	15000
GGGATTTTGT	TCAGGGATTT	ATAAGTTATG	TAGGGTCTCA	TTATAGAGAA	GTGAACCTTG	15060
TATACCCTAG	ATACAGCAAC	TTCATCTCTA	CTGAATCTTA	TTTGGTTATG	ACAGATCTCA	15120
AGGCTAACCG	GCTAATGAAT	CCTGAAAAGA	TTAAGCAGCA	GATAATTGAA	TCATCTGTGA	15180
GGACTTCACC	TGGACTTATA	GGTCACATCC	TATCCATTAA	GCAACTAAGC	TGCATACAAG	15240
CAATTGTGGG	AGACGCGATT	AGTAGAGGTG	ATATCAATCC	TACTCTGAAA	AACTTACAC	15300
CTATAGAGCA	GGTGCTGATC	AATTGCGGGT	TGGCAATTAA	CGGACCTAAG	CTGTGCAAAG	15360
AATTGATCCA	CCATGATGTT	GCCTCAGGGC	AAGATGGATT	GCTTAATTCT	ATACTCATCC	15420
TCTACAGGGA	GTTGGCAAGA	TTCAAAGACA	ACCAAAGAAG	TCAACAAGGG	ATGTTCCACG	15480
CTTACCCCGT	ATTGGTAAGT	AGCAGGCAAC	GAGAACTTAT	ATCTAGGATC	ACCCGCAAAT	15540
TCTGGGGGCA	CATTCTTCTT	TACTCCGGGA	ACAGAAAGTT	GATAAATAAG	TTTATCCAGA	15600
ATCTCAAGTC	CGGCTATCTG	ATACTAGACT	TACACCAGAA	TATCTTCGTT	AAGAATCTAT	15660
CCAAGTCAGA	GAAACAGATT	ATTATGACGG	GGGGTTTGAA	ACGTGAGTGG	GTTTTTAAGG	15720
TAACAGTCAA	GGAGACCAA	GAATGGTATA	AGTTAGTCGG	ATACAGTGCC	CTGATTAAGG	15780
ACTAATTGGT	TGAACTCCGG	AACCCTAATC	CTGCCCTAGG	TGGTTAGGCA	TTATTTGCAA	15840

SUBSTITUTE SHEET (RULE 26)

- 172 -

TATATTAAAG AAAACTTTGA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT

15894

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu
1           5           10           15

Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala
          20           25           30

Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn
          35           40           45

Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn
          50           55           60

Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro
          65           70           75           80

Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn
          85           90           95

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys
          100          105          110

Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu
          115          120          125

Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp
          130          135          140

Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln
          145          150          155          160

Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg
          165          170          175

Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr

```

SUBSTITUTE SHEET (RULE 26)

- 173 -

180	185	190
Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp 195 200 205		
Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr 210 215 220		
Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met 225 230 235 240		
Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly 245 250 255		
Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu 260 265 270		
Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu 275 280 285		
Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe 290 295 300		
Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly 305 310 315 320		
Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr 325 330 335		
Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe 340 345 350		
Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu 355 360 365		
Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr 370 375 380		
Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr 385 390 395 400		
Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His 405 410 415		
Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr 420 425 430		
His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe 435 440 445		
Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu 450 455 460		

SUBSTITUTE SHEET (RULE 26)

- 174 -

Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr  
 465 470 475 480  
 Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg  
 485 490 495  
 Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp  
 500 505 510  
 Val Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe  
 515 520 525  
 Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg  
 530 535 540  
 Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala  
 545 550 555 560  
 Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly  
 565 570 575  
 Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala  
 580 585 590  
 Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro  
 595 600 605  
 Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn  
 610 615 620  
 Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Val Ile Arg Gln  
 625 630 635 640  
 Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val  
 645 650 655  
 Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670  
 Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly  
 675 680 685  
 Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val  
 690 695 700  
 Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile  
 705 710 715 720  
 Pro Leu Tyr Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met  
 725 730 735

SUBSTITUTE SHEET (RULE 26)

- 175 -

Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile  
 740 745 750  
 Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser  
 755 760 765  
 Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro  
 770 775 780  
 Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr  
 785 790 795 800  
 Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His  
 805 810 815  
 His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr  
 820 825 830  
 Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845  
 Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860  
 Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880  
 Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895  
 Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910  
 Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile  
 915 920 925  
 Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940  
 Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960  
 Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu  
 965 970 975  
 Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990  
 Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005  
 Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His

SUBSTITUTE SHEET (RULE 26)

- 176 -

1010	1015	1020
Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu 1025	1030	1035 1040
Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val 1045	1050	1055
Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg 1060	1065	1070
Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala 1075	1080	1085
Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser 1090	1095	1100
Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly 1105	1110	1115 1120
Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu 1125	1130	1135
Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg 1140	1145	1150
Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly 1155	1160	1165
His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser 1170	1175	1180
Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp 1185	1190	1195 1200
Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr 1205	1210	1215
Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser 1220	1225	1230
Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala 1235	1240	1245
Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg 1250	1255	1260
Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile 1265	1270	1275 1280
Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln 1285	1290	1295

SUBSTITUTE SHEET (RULE 26)



- 177 -

Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr  
1300 1305 1310

Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp  
1315 1320 1325

Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu  
1330 1335 1340

Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val  
1345 1350 1355 1360

Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp  
1365 1370 1375

His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu  
1380 1385 1390

Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp  
1395 1400 1405

Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe  
1410 1415 1420

Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr  
1425 1430 1435 1440

Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met  
1445 1450 1455

Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile  
1460 1465 1470

Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly  
1475 1480 1485

Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro  
1490 1495 1500

Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg  
1505 1510 1515 1520

Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro  
1525 1530 1535

Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His  
1540 1545 1550

Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met  
1555 1560 1565

SUBSTITUTE SHEET (RULE 26)

- 178 -

Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu  
 1570 1575 1580  
 Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val  
 1585 1590 1595 1600  
 Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala  
 1605 1610 1615  
 Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg  
 1620 1625 1630  
 Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala  
 1635 1640 1645  
 Met Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val  
 1650 1655 1660  
 Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
 1665 1670 1675 1680  
 Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
 1685 1690 1695  
 Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
 1700 1705 1710  
 Met Ser Ile Lys Ala Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
 1715 1720 1725  
 Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
 1730 1735 1740  
 Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn  
 1745 1750 1755 1760  
 Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
 1765 1770 1775  
 Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
 1780 1785 1790  
 Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
 1795 1800 1805  
 Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
 1810 1815 1820  
 Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
 1825 1830 1835 1840  
 Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp

SUBSTITUTE SHEET (RULE 26)

- 179 -

1845	1850	1855
Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr		
1860	1865	1870
Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asp Lys		
1875	1880	1885
Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala		
1890	1895	1900
Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro		
1905	1910	1915
Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His		
1925	1930	1935
Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser		
1940	1945	1950
Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met		
1955	1960	1965
Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr		
1970	1975	1980
Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys		
1985	1990	1995
Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro		
2005	2010	2015
Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly		
2020	2025	2030
Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp		
2035	2040	2045
Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr		
2050	2055	2060
Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met		
2065	2070	2075
Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile		
2085	2090	2095
Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly		
2100	2105	2110
Asn Arg Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr		
2115	2120	2125

SUBSTITUTE SHEET (RULE 26)

- 180 -

Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
 2130 2135 2140

Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
 2145 2150 2155 2160

Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
 2165 2170 2175

Tyr Ser Ala Leu Ile Lys Asp  
 2180

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACCAAACAAA GTTGGGTAAG GATAGTTCAA TCAATGATCA TCTTCTAGTG CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCTCTCA	240
TTACCACTCG ATCCAGACTT CTGGACCGGT TGGTGAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAATTAG TAGTGATCAA TCCAGGTTCTG	540
GATGGTTCGG GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCCCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCT GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720

SUBSTITUTE SHEET (RULE 26)

- 181 -

TAGTTGGTGA	ATTAGATTG	GAGAGAAAAT	GGTTGGATGT	GGTGAGGAAC	AGGATTGCCG	780
AGGACCTCTC	CTTACGCCGA	TTCATGGTCG	CTCTAATCCT	GGATATCAAG	AGAACACCCG	840
GAAACAAACC	CAGGATTGCT	GAAATGATAT	GTGACATTGA	TACATATATC	GTAGAGGCAG	900
GATTAGCCAG	TTTTATCCTG	ACTATTAAGT	TTGGGATAGA	AACTATGTAT	CCTGCTCTTG	960
GACTGCATGA	ATTTGCTGGT	GAGTTATCCA	CACTTGAGTC	CTTGATGAAC	CTTTACCAGC	1020
AAATGGGGGA	AACTGCACCC	TACATGGTAA	TCCTGGAGAA	CTCAATTCAG	AACAAGTTCA	1080
GTGCAGGATC	ATACCCTCTG	CTCTGGAGCT	ATGCCATGGG	AGTAGGAGTG	GAACCTGAAA	1140
ACTCCATGGG	AGGTTTGAAC	TTTGGCCGAT	CTTACTTTGA	TCCAGCATAT	TTTAGATTAG	1200
GGCAAGAGAT	GGTAAGGAGG	TCAGCTGGAA	AGGTCAGTTC	CACATTGGCA	TCTGAACTCG	1260
GTATCACTGC	CGAGGATGCA	AGGCTTGTTT	CAGAGATTGC	AATGCATACT	ACTGAGGACA	1320
AGATCAGTAG	AGCGGTTGGA	CCCAGACAAG	CCCAAGTATC	ATTTCTACAC	GGTGATCAAA	1380
GTGAGAATGA	GCTACCGAGA	TTGGGGGGCA	AGGAAGATAG	GAGGGTCAAA	CAGAGTCGAG	1440
GAGAAGCCAG	GGAGAGCTAC	AGAGAAACCG	GGCCCAGCAG	AGCAAGTGAT	GCGAGAGCTG	1500
CCCATCTTCC	AACCGGCACA	CCCCTAGACA	TTGACACTGC	AACGGAGTCC	AGCCAAGATC	1560
CGCAGGACAG	TCGAAGGTCA	GCTGACGCCC	TGCTTAGGCT	GCAAGCCATG	GCAGGAATCT	1620
CGGAAGAACA	AGGCTCAGAC	ACGGACACCC	CTATAGTGTA	CAATGACAGA	AATCTTCTAG	1680
ACTAGGTGCG	AGAGGCCGAG	GGCCAGAACA	ACATCCGCCT	ACCATCCATC	ATTGTTATAA	1740
AAAACCTTAGG	AACCAGGTCC	ACACAGCCGC	CAGCCCATCA	ACCATCCACT	CCCACGATTG	1800
GAGCCAATGG	CAGAAGAGCA	GGCACGCCAT	GTCAAAAACG	GACTGGAATG	CATCCGGGCT	1860
CTCAAGGCCG	AGCCCATCGG	CTCACTGGCC	ATCGAGGAAG	CTATGGCAGC	ATGGTCAGAA	1920
ATATCAGACA	ACCCAGGACA	GGAGCGAGCC	ACCTGCAGGG	AAGAGAAGGC	AGGCAGTTCC	1980
GGTCTCAGCA	AACCATGCCT	CTCAGCAATT	GGATCAACTG	AAGGCGGTGC	ACCTCGCATC	2040
CGCGGTCAAG	GACCTGGAGA	GAGCGATGAC	GACGCTGAAA	CTTTGGGAAT	CCCCCAAGA	2100
AATCTCCAGG	CATCAAGCAC	TGGGTTACAG	TGTTATTACG	TTTATGATCA	CAGCGGTGAA	2160
GCGGTTAAGG	GAATCCAAGA	TGCTGACTCT	ATCATGGTTC	AATCAGGCCT	TGATGGTGAT	2220
AGCACCTCT	CAGGAGGAGA	CAATGAATCT	GAAAACAGCG	ATGTGGATAT	TGGCGAACCT	2280

SUBSTITUTE SHEET (RULE 26)

- 182 -

GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTCAGG 2340  
GCTTCTGATG TTGAAACTGC AGAAGGAGGG GAGATCCACG AGCTCCTGAG ACTCCAATCC 2400  
AGAGGCAACA ACTTTCGGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC CCCGGACCCC 2460  
GGTAGGGCCA GCACTTCCGG GACACCCATT AAAAAGGGCA CAGACGCGAG ATTAGCCTCA 2520  
TTTGGAACGG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA 2580  
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GCGGGGAATG TCCCCGAGTG TGTGAGCAAT 2640  
GCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCCAG 2700  
AATAATGAAG AAGGGGGAGA CTATTATGAT GATGAGCTGT TCTCTGATGT CCAAGATATT 2760  
AAAACAGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCTCCAA GCTAGAATCA 2820  
CTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAGAAGC AGATCAACAG GCAAATATC 2880  
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG 2940  
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA 3000  
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCGTTGC CAGCCGACAA 3060  
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTGAG 3120  
CTAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCTGA CACCGGCCCT 3180  
GCATCACGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG 3240  
CGTTACCTGA TGAATCTCCT TGATGATATC AAAGGAGCCA ATGATCTTGC CAAGTTCCAC 3300  
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG 3360  
CCAGTCGACC CAACTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT 3420  
GCCTCCCAAG GTCCACAATG ACAGAGACCT ACGACTTCGA CAAGTCGGCA TGGGACATCA 3480  
AAGGGTCGAT CGCTCCGATA CAACCCACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG 3540  
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTATG TACATGTTTC 3600  
TGCTGGGGGT TGTTGAGGAC AGCGATTCCC TAGGGCCTCC AATCGGGCGA GCATTGGGT 3660  
TCCTGCCCTT AGGTGTTGGC AGATCCACAG CAAAGCCCGA AAAACTCCTC AAAGAGGCCA 3720  
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA 3780  
ACAACACCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT 3840

SUBSTITUTE SHEET (RULE 26)

- 183 -

TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTCGATACC CCGCAGAGGT	3900
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTCTTA	3960
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCTTA	4020
GGATTGACAA GGCATAGGC CCTGGGAAGA TCATCGACAA TACAGAGCAA CTTCCTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCACTTGGT GGGATAGGGG	4200
GCACCACTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACTCG	4260
GGTTCAAGAA GACCTTATGT TACCCGCTGA TGGATATCAA TGAAGACCTT AATCGATTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC	4380
AAGAATTCCG CATTTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC	4440
TGTAGACCGT AGTGCCCGAGC AATGCCCCGAA AACGACCCCC CTCACAATGA CAGCCAGAAG	4500
GCCCCGACAA AAAAGCCCCC TCCGAAAGAC TCCACGGACC AAGCGAGAGG CCAGCCAGCA	4560
GCCGACGGCA AGCGCGAACA CCAGGCGGCC CCAGCACAGA ACAGCCCTGA CACAAGGCCA	4620
CCACCAGCCA CCCCAATCTG CATCCTCCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGC	4680
TGCCCCGAT CCAAACCACC AACCGCATCC CCACCACCCC CGGGAAAGAA ACCCCCAGCA	4740
ATTGGAAGGC CCCTCCCCCT CTTCTCAAC ACAAGAACTC CACAACCGAA CCGCACAAGC	4800
GACCGAGGTG ACCCAACCGC AGGCATCCGA CTCCCTAGAC AGATCCTCTC TCCCCGGCAA	4860
ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAC AGAACCAGCA CCCCAGTCCA	4920
CGGTGCCGCG CCCCCAACCC CCGACAACCA GAGGGAGCCC CCAACCAATC CCGCCGGCTC	4980
CCCCGGTGCC CACAGGCAGG GACACCAACC CCCGAACAGA CCCAGCACCC AACCATCGAC	5040
AATCCAAGAC GGGGGGGCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCACCGC	5100
GAGGAAGCCC ACCCACCCTA CACACGACCA CGGCAACCAA ACCAGAACCC AGACCACCCT	5160
GGGCCACCAG CTCCCAGACT CGGCCATCAC CCCGCAGAAA GGAAAGGCCA CAACCCGCGC	5220
ACCCAGCCC CGATCCGGCG GGGAGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC	5280
CGAAGGACCC CCGAACCGCA AAGGACATCA GTATCCACA GCCTCTCAA GTCCCCGGT	5340
CTCCTCCTCT TCTGAAGGG ACCAAAAGAT CAATCCACCA CACCCGACGA CACTCAACTC	5400

SUBSTITUTE SHEET (RULE 26)

- 184 -

CCCACCCCTA	AAGGAGACAC	CGGGAATCCC	AGAATCAAGA	CTCATCCAAT	GTCCATCATG	5460
GGTCTCAAGG	TGAACGTCTC	TGCCATATTC	ATGGCAGTAC	TGTTAACTCT	CCAAACACCC	5520
ACCGGTCAAA	TCCATTGGGG	CAATCTCTCT	AAGATAGGGG	TGGTAGGAAT	AGGAAGTGCA	5580
AGCTACAAAG	TTATGACTCG	TTCCAGCCAT	CAATCATTAG	TCATAAAATT	AATGCCCAAT	5640
ATAACTCTCC	TCAATAACTG	CACGAGGGTA	GAGATTGCAG	AATACAGGAG	ACTACTGAGA	5700
ACAGTTTTTG	AACCAATTAG	AGATGCACTT	AATGCAATGA	CCCAGAATAT	AAGACCGGTT	5760
CAGAGTGTAG	CTTCAAGTAG	GAGACACAAG	AGATTTGCGG	GAGTAGTCCT	GGCAGGTGCG	5820
GCCCTAGGCG	TTGCCACAGC	TGCTCAGATA	ACAGCCGGCA	TTGCACTTCA	CCAGTCCATG	5880
CTGAACTCTC	AAGCCATCGA	CAATCTGAGA	GCGAGCCTGG	AAACTACTAA	TCAGGCAATT	5940
GAGACAATCA	GACAAGCAGG	GCAGGAGATG	ATATTGGCTG	TTCAGGGTGT	CCAAGACTAC	6000
ATCAATAATG	AGCTGATACC	GTCTATGAAC	CAACTATCTT	GTGATTTAAT	CGGCCAGAAG	6060
CTCGGGCTCA	AATTGCTCAG	ATACTATACA	GAAATCCTGT	CATTATTTGG	CCCCAGTTTA	6120
CGGGACCCCA	TATCTGCGGA	GATATCTATC	CAGGCTTTGA	GCTATGCGCT	TGGAGGAGAC	6180
ATCAATAAGG	TGTTAGAAAA	GCTCGGATAC	AGTGGAGGTG	ATTTACTGGG	CATCTTAGAG	6240
AGCGGAGGAA	TAAAGGCCCG	GATAACTCAC	GTCGACACAG	AGTCCTACTT	CATTGTCCTC	6300
AGTATAGCCT	ATCCGACGCT	GTCCGAGATT	AAGGGGGTGA	TTGTCCACCG	GCTAGAGGGG	6360
GTCTCGTACA	ACATAGGCTC	TCAAGAGTGG	TATACCACTG	TGCCCAAGTA	TGTTGCAACC	6420
CAAGGGTACC	TTATCTCGAA	TTTTGATGAG	TCATCGTGTA	CTTTCATGCC	AGAGGGGACT	6480
GTGTGCAGCC	AAAATGCCTT	GTACCCGATG	AGTCCTCTGC	TCCAAGAATG	CCTCCGGGGG	6540
TACACCAAGT	CCTGTGCTCG	TACACTCGTA	TCCGGGTCTT	TTGGGAACCG	GTTTCATTTA	6600
TCACAAGGGA	ACCTAATAGC	CAATTGTGCA	TCAATCCTTT	GCAAGTGTTA	CACAACAGGA	6660
ACGATCATT	ATCAAGACCC	TGACAAGATC	CTAACATACA	TTGCTGCCGA	TCACTGCCCC	6720
GTAGTCGAGG	TGAACGGCGT	GACCATCCAA	GTCGGGAGCA	GGAGGTATCC	AGACGCTGTG	6780
TACTTGACCA	GAATTGACCT	CGGTCCTCCC	ATATCATTGG	AGAGGTTGGA	CGTAGGGACA	6840
AATCTGGGGA	ATGCAATTGC	TAAGTTGGAG	GATGCCAAGG	AATTGTTGGA	GTCATCGGAC	6900
CAGATATTGA	GGAGTATGAA	AGGTTTATCG	AGCACTAGCA	TAGTCTACAT	CCTGATTGCA	6960

SUBSTITUTE SHEET (RULE 26)



- 185 -

GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACGGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCACCCA GCATCAAGCC CACCTGAAAT	7200
TATCTCCGGC TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATCAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA	7320
TCCCAAGGGA AGTAGGATAG TCATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTGTTTG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500
TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TCACTGACC TAGTGAAATT	7620
AATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCCTGTTAGA CTTGTATTTA GGTCGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTATGGGGGA ACTTACCTAG TGGAAAAGCC	7980
TAATCTGAGC AGCAAAAGGT CAGAGTTGTC ACAACTGAGC ATGTACCGAG TGTTTGAAGT	8040
AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATCTTGA	8100
GCAACCAGTC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAACCTCGC	8160
AGCCCTTTGT CACGGGGAAG ATTCTATCAC AATTCCTTAT CAGGGATCAG GGAAAGGTGT	8220
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CCCCTTATCA ACGGATGATC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGAACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAAC AGGCGTGTA GGGTAAATC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTGATCT	8520

SUBSTITUTE SHEET (RULE 26)

- 186 -

GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA	8580
CGGTTCAGGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGACTATCCC	8640
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTTAGTCCC TACCTCTTCA CTGTCCCAAT TAAGGAAGCA GGCGAAGACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAATC AGTTCCAATC TGGTGATTCT	8820
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCCAGGG TTGAACATGC	8880
TGTGGTTTAT TACGTTTACA GCCCAAGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT	8940
GCCTATAAAG GGGGTCCCCA TCGAATTACA AGTGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC	9060
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATCGCAG	9120
ATAGGGCTGC TAGTGAACCA ATCACATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACCGCC TAAAAACGG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCACT	9480
CTCATATTCC ATATCCAAAT TGAATCAGG ATTTATTAA CATAGAAGAC AAAGAGTCAA	9540
CGAGGAAGAT CCGTGAATC CTCAAAAAGG GGAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCTTAAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAAGTTGG GAGTTTACAT GCACAGCTCC CAGTGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACATTGA ACTGGTTTGG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080

SUBSTITUTE SHEET (RULE 26)

- 187 -

TGGAGCCTCT TTCACTTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG 10140  
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG 10200  
ATGAAGGTAC TTATCATGAG TTAAGTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA 10260  
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG 10320  
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG 10380  
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA 10440  
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA 10500  
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAATCTT 10560  
TTGCTGGAGT GAAATTTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT 10620  
ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG 10680  
AGTTCCTGCG TTACGACCCT CCCAAGGGAA CCGGGTCACG GAGGCTTGTA GATGTTTTCC 10740  
TTAATGATTC GAGCTTTGAC CCATATGATG TGATAATGTA TGTTGTAAGT GGAGCTTACC 10800  
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG 10860  
GTAGACTTTT TGCTAAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC 10920  
TAATCTCAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCAG 10980  
ATTTGACTAA GGCACTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA 11040  
GTCACAGGGG GGGGCCAGTC TTA AAAACCT ACTCCGAAG CCCAGTCCAC ACAAGTACCA 11100  
GGAACGTGAG AGCAGCAAAA GGGTTTATAG GGTTCCTCA AGTAATTCGG CAGGACCAAG 11160  
ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATT ATCACGACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTG CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCTGTGA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCCAT ATCCCGTTAT 11400  
ATAAAGTCCC CAATGATCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATTC CCTATCTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCCTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640

SUBSTITUTE SHEET (RULE 26)

- 188 -

ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CTTGTCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTCTGATCT 11940  
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCCTCACTAA TGCCTGAAGA GACCCTCCAT CAAGTAATGA 12180  
CACAACAACC GGGGGACTCT TCATTCCCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240  
TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA 12300  
TCCATAGTCC AAACCCAATG TTAAAAGGAT TATTCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC GGCATTCCCTC ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCA CA GGGGCAAGAG AGTCTATTGC AGGCATGCTG GATACCACAA 12480  
AAGGCTTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGA CTATGAA CAATTCAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GGCGAGAGCT CTAAGAAGCC 12660  
ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC 12720  
TAGAATCTAT GCGAGGCCAC CTTATTCGGC GTCATGAGAC ATGTGTCATC TGCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTTTGTCC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTGCTAAGA GCCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA 12960  
CAGTGTA CTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG 13020  
CTAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCGCATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTC CG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT 13200

SUBSTITUTE SHEET (RULE 26)

- 189 -

CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAAGG AATGCTTCTA GGGTTGGGTG 13260  
TTTTAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TCGTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CCCGCAAGCT AGAGCTGAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATGCA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG 13500  
AATTTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTCATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740  
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTGTGCATC GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGCACA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTCTTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA 14040  
GATTCGACAA CATCCAGGCA AAACACTTAT GTGTTCTGGC AGATTTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCAATTCGA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC 14160  
ATATCAAGGC AGAGGCTATG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTACTIONATG TCTCTGACTT ATCTCCGGCG AGGATCGATC AAACAGATAA 14280  
GATTGAGAGT TGATCCAGGA TTCATTTTCG ACGCCCTCGC TGAGGTAAAT GTCAGTCAGC 14340  
CAAAGATCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGCTTTCAGA CCCCCACACG 14400  
ATGATGTTGC AAAATTGCTC AAAGATATCA ACACAAGCAA GCACAATCTT CCCATTTTCAG 14460  
GGGGCAATCT CGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTG AACTCATCTG 14520  
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAGG 14580  
ACGGCTTGTT CTTGGGTGAG GGATCGGGTT CTATGTTGAT CACTTATAAA GAGATACTTA 14640  
AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG 14700  
AATTAGCACC CTATCCCTCC GAAGTTGGCC TTGTCGAACA CAGAATGGGA GTAGGTAATA 14760

SUBSTITUTE SHEET (RULE 26)

- 190 -

TTGTCAAAGT GCTCTTTAAC GGGAGGCCCG AAGTCACGTG GGTAGGCAGT GTAGATTGCT 14820  
TCAATTTTCAT AGTTAGTAAT ATCCCTACCT CTAGTGTGGG GTTTATCCAT TCAGATATAG 14880  
AGACCTTGCC TGACAAAGAT ACTATAGAGA AGCTAGAGGA ATTGGCAGCC ATCTTATCGA 14940  
TGGCTCTGCT CCTGGGCAAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTCAGCG 15000  
GGGATTTTGT TCAGGGATT ATAAGTTATG TAGGGTCTCA TTATAGAGAA GTGAACCTTG 15060  
TATACCCTAG ATACAGCAAC TTCATCTCTA CTGAATCTTA TTTGGTTATG ACAGATCTCA 15120  
AGGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGA 15180  
GGACTTCACC TGGACTTATA GGTCCATCC TATCCATTAA GCAACTAAGC TGCATACAAG 15240  
CAATTGTGGG AGACGCAGTT AGTAGAGGTG ATATCAATCC TACTCTGAAA AAACCTACAC 15300  
CTATAGAGCA GGTGCTGATC AATTGCGGGT TGGCAATTAA CGGACCTAAG CTGTGCAAAG 15360  
AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAATTCT ATACTCATCC 15420  
TCTACAGGGA GTTGCAAGA TTCAAAGACA ACCAAAGAAG TCAACAAGGG ATGTTCCACG 15480  
CTTACCCCGT ATTGGTAAGT AGCAGGCAAC GAGAACTTAT ATCTAGGATC ACCCGCAAAT 15540  
TCTGGGGGCA CATTCTTCTT TACTCCGGGA ACAAAAAGTT GATAAATAAG TTTATCCAGA 15600  
ATCTCAAGTC CGGCTATCTG ATACTAGACT TACACCAGAA TATCTTCGTT AAGAATCTAT 15660  
CCAAGTCAGA GAAACAGATT ATTATGACGG GGGGTTTGAA ACGTGAGTGG GTTTTAAAGG 15720  
TAACAGTCAA GGAGACCAAA GAATGGTATA AGTTAGTCGG ATACAGTGCC CTGATTAAGG 15780  
ACTAATTGGT TGAATCCGG AACCTAATC CTGCCTAGG TGGTAGGCA TTATTTGCAA 15840  
TATATTAAAG AAACTTTGA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT 15894

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

SUBSTITUTE SHEET (RULE 26)

- 191 -

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
1 5 10 15

Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
20 25 30

Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
35 40 45

Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
50 55 60

Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
65 70 75 80

Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn  
85 90 95

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys  
100 105 110

Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu  
115 120 125

Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp  
130 135 140

Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln  
145 150 155 160

Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg  
165 170 175

Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr  
180 185 190

Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
195 200 205

Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
210 215 220

Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
225 230 235 240

Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly  
245 250 255

Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
260 265 270

SUBSTITUTE SHEET (RULE 26)

- 192 -

Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu  
 275 280 285  
 Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe  
 290 295 300  
 Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly  
 305 310 315 320  
 Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Thr Glu Ala Leu Asp Tyr  
 325 330 335  
 Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe  
 340 345 350  
 Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu  
 355 360 365  
 Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr  
 370 375 380  
 Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr  
 385 390 395 400  
 Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His  
 405 410 415  
 Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr  
 420 425 430  
 His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe  
 435 440 445  
 Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu  
 450 455 460  
 Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr  
 465 470 475 480  
 Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg  
 485 490 495  
 Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp  
 500 505 510  
 Val Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe  
 515 520 525  
 Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg  
 530 535 540  
 Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala

SUBSTITUTE SHEET (RULE 26)



- 193 -

545	550	555	560
Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly	565	570	575
Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala	580	585	590
Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro	595	600	605
Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn	610	615	620
Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Val Ile Arg Gln	625	630	635
Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val	645	650	655
Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg	660	665	670
Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly	675	680	685
Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val	690	695	700
Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile	705	710	715
Pro Leu Tyr Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met	725	730	735
Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile	740	745	750
Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser	755	760	765
Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro	770	775	780
Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr	785	790	795
Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His	805	810	815
His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr	820	825	830

SUBSTITUTE SHEET (RULE 26)

- 194 -

Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845

Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860

Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880

Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895

Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910

Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile  
 915 920 925

Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940

Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960

Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu  
 965 970 975

Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990

Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005

Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His  
 1010 1015 1020

Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
 1025 1030 1035 1040

Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
 1045 1050 1055

Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
 1060 1065 1070

Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
 1075 1080 1085

Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
 1090 1095 1100

SUBSTITUTE SHEET (RULE 26)

- 195 -

Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly  
 1105 1110 1115 1120

Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu  
 1125 1130 1135

Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg  
 1140 1145 1150

Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly  
 1155 1160 1165

His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser  
 1170 1175 1180

Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp  
 1185 1190 1195 1200

Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr  
 1205 1210 1215

Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser  
 1220 1225 1230

Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala  
 1235 1240 1245

Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg  
 1250 1255 1260

Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile  
 1265 1270 1275 1280

Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln  
 1285 1290 1295

Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr  
 1300 1305 1310

Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp  
 1315 1320 1325

Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu  
 1330 1335 1340

Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val  
 1345 1350 1355 1360

Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp  
 1365 1370 1375

His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu

SUBSTITUTE SHEET (RULE 26)

- 196 -

1380	1385	1390
Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp 1395	1400	1405
Ala Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe 1410	1415	1420
Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr 1425	1430	1435 1440
Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met 1445	1450	1455
Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile 1460	1465	1470
Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly 1475	1480	1485
Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro 1490	1495	1500
Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg 1505	1510	1515 1520
Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro 1525	1530	1535
Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His 1540	1545	1550
Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met 1555	1560	1565
Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu 1570	1575	1580
Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val 1585	1590	1595 1600
Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala 1605	1610	1615
Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg 1620	1625	1630
Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala 1635	1640	1645
Met Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val 1650	1655	1660

SUBSTITUTE SHEET (RULE 26)

- 197 -

Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
1665 1670 1675 1680

Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
1685 1690 1695

Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
1700 1705 1710

Met Ser Ile Lys Ala Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
1715 1720 1725

Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
1730 1735 1740

Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn  
1745 1750 1755 1760

Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
1765 1770 1775

Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
1780 1785 1790

Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
1795 1800 1805

Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
1810 1815 1820

Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
1825 1830 1835 1840

Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp  
1845 1850 1855

Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr  
1860 1865 1870

Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asp Lys  
1875 1880 1885

Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
1890 1895 1900

Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
1905 1910 1915 1920

Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His  
1925 1930 1935

SUBSTITUTE SHEET (RULE 26)

- 198 -

Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser  
1940 1945 1950

Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met  
1955 1960 1965

Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr  
1970 1975 1980

Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys  
1985 1990 1995 2000

Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro  
2005 2010 2015

Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly  
2020 2025 2030

Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp  
2035 2040 2045

Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr  
2050 2055 2060

Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met  
2065 2070 2075 2080

Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile  
2085 2090 2095

Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly  
2100 2105 2110

Asn Lys Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr  
2115 2120 2125

Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
2130 2135 2140

Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
2145 2150 2155 2160

Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
2165 2170 2175

Tyr Ser Ala Leu Ile Lys Asp  
2180

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15894 base pairs

SUBSTITUTE SHEET (RULE 26)

- 199 -

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACCAAACAAA GTTGGGTAAG GATAGTTCAA TCAATGATCA TTTTCTAGTG CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAAGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180
GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCCTCAA	240
TTACCACTCG ATCCAGACTT CTGGACCGGT TGGTCAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CCAATACTTT TCACATGATG ATCCAATTAG TAGTGATCAA TCCAGGTTCTG	540
GATGGTTCGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCTCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGAGTGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGAACACCCG	840
GAAACAAACC CAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GACTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAC CTTTACCAGC	1020
AAATGGGGGA AACTGCACCC TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAGTTCA	1080
GTGCAGGATC ATACCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGCCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTAAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCTGAACTCG	1260

SUBSTITUTE SHEET (RULE 26)

- 200 -

GTATCACTGC CGAGGATGCA AGGCTTGTTT CAGAGATTGC AATGCATACT ACTGAGGACA	1320
AGATCAGTAG AGCGGTTGGA CCCAGACAAG CCCAAGTATC ATTTCTACAC GGTGATCAAA	1380
GTGAGAATGA GCTACCGAGA TTGGGGGGCA AGGAAGATAG GAGGGTCAAA CAGAGTCGAG	1440
GAGAAGCCAG GGAGAGCTAC AGAGAAACCG GGCCAGCAG AGCAAGTGAT GCGAGAGCTG	1500
CCCATCTTCC AACCGGCACA CCCCTAGACA TTGACACTGC ATCGGAGTCC AGCCAAGATC	1560
CGCAGGACAG TCGAAGGTCA GCTGACGCCC TGCTTAGGCT GCAAGCCATG GCAGGAATCT	1620
CGGAAGAACA AGGCTCAGAC ACGGACACCC CTATAGTGTA CAATGACAGA AATCTTCTAG	1680
ACTAGGTGCG AGAGGCCGAG GGCCAGAACA ACATCCGCTT ACCCTCCATC ATTGTTATAA	1740
AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCCATCA ACCATCCACT CCCACGATTG	1800
GAGCCGATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGCT	1860
CTCAAGGCCG AGCCCATCGG CTCCTGGCC ATCGAGGAAG CTATGGCAGC ATGGTCAGAA	1920
ATATCAGACA ACCCAGGACA GGAGCGAGCC ACCTGCAGGG AAGAGAAGGC AGGCAGTTCC	1980
GGTCTCAGCA AACCATGCCT CTCAGCAATT GGATCAACTG AAGGCGGTGC ACCTCGCATC	2040
CGCGGTCAGG GACCTGGAGA GAGCGATGAC GACGCTGAAA CTTTGGGAAT CCCCCAAGA	2100
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTATTATG TTTATGATCA CAGCGGTGAA	2160
GCGGTTAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT	2220
AGCACCCCTCT CAGGAGGAGA CAATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT	2280
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTCCAGG	2340
GCTTCTGATG TTGAAACTGC AGAAGGAGGG GAGATCCACG AGCTCCTGAG ACTCCAATCC	2400
AGAGGCAACA ACTTTCCGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC TCCGGACCCC	2460
GGTAGGGCCA GCACTTCCGG GACACCCATT AAAAAAGGCA CAGACGCGAG ATTAGCCTCA	2520
TTTGGAACGG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA	2580
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GCGGGGAATG TCCCCGAGTG TGTGAGCAAT	2640
GCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCCAG	2700
AATAATGAAG AAGGGGGAGA CTATTATGAT GATGAGCTGT TCTCTGATGT CCAAGATATT	2760
AAAACAGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCTCCAA GCTAGAATCA	2820

SUBSTITUTE SHEET (RULE 26)



- 201 -

CTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAGAAGC AGATCAACAG GCAAAATATC	2880
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG	2940
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA	3000
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCGTTGC CAGCCGACAA	3060
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTTCA	3120
CTAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCTTGA CACCGGCCCT	3180
GCATCACGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG	3240
CGTTACCTGA TGACTCTCCT TGATGATATC AAAGGAGCCA ATGATCTTGC CAAGTTCCAC	3300
CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG	3360
CCAGTCGACC CAACTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT	3420
GCCTCCCAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA	3480
AAGGGTTGAT CGCTCCGATA CAACCCACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG	3540
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTATG TACATGTTTC	3600
TGCTGGGGGT TGTGAGGAC AGCGATCCCC TAGGGCCTCC AATCGGGCGA GCATTTGGGT	3660
CCCTGCCCTT AGGTGTTGGC AAATCCACAG CAAAGCCCGA AAAACTCCTC AAAGAGGCCA	3720
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA	3780
ACAAACCCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT	3840
TCAACGCAAA CCAAGTGTGC AGTGCGGTTA ATCTGATACC GCTCGATACC CCGCAGAGGT	3900
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTCTTA	3960
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCTTA	4020
GGATTGACAA GGCGATAGGC CCTGGGAAGA TCATCGACAA TACAGAGCAA CTTCTTGAGG	4080
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG	4140
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCATTGGT GGGATAGGGG	4200
GCACCAGTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACTCG	4260
GGTTCAAGAA GACCTTATGT TACCCGCTGA TAGATATCAA TGAAGACCTT AATCGATTAC	4320
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC	4380

SUBSTITUTE SHEET (RULE 26)

- 202 -

AAGAATTCCG	CATTTACGAC	GACGTGATCA	TAAATGATGA	CCAAGGACTA	TTCAAAGTTC	4440
TGTAGACCGT	AGTGCCCGAGC	AATGCCCCGAA	AACGACCCCC	CTCACAATGA	CAGCCAGAAG	4500
GCCCCGGACAA	AAAAGCCCCC	TCCGAAAGAC	TCCACGGACC	AAGCGAGAGG	CCAGCCAGCA	4560
GCCGACGGCA	AGCGCGAACA	CCAGGCGGCC	CCAGCACAGA	ACAGCCCTGA	TACAAGGCCA	4620
CCACCAGCCA	CCCCAATCTG	CATCCTCCTC	GTGGGACCCC	CGAGGACCAA	CCCCCAAGGC	4680
TGCCCCCGAT	CCAAACCACC	AACCGCATCC	CCACCACCCC	CGGGAAAGAA	ACCCCCAGCA	4740
ATTGGAAGGC	CCCTCCCCCT	CTTCTCAAC	ACAAGAACTC	CACAACCGAA	CCGCACAAGC	4800
GACCGAGGTG	ACCCAACCGC	AGGCATCCGA	CTCCCTAGAC	AGATCCTCTC	TCCCCGGCAA	4860
ACTAAACAAA	ACTTAGGGCC	AAGGAACATA	CACACCCAAC	AGAACCCAGA	CCCCGGCCCA	4920
CGGCGCCGCG	CCCCCAACCC	CCGACAACCA	GAGGGAGCCC	CCAACCAATC	CCGCCGGCTC	4980
CCCCGGTGCC	CACAGGCAGG	GACACCAACC	CCCGAACAGA	CCCAGCACCC	AACCATCGAC	5040
AATCCAAGAC	GGGGGGGGCC	CCCCAAAAAA	AGGCCCCCAG	GGGCCGACAG	CCAGCACC GC	5100
GAGGAAGCCC	ACCCACCCCA	CACACGACCA	CGGCAACCAA	ACCAGAACCC	AGACCACCTT	5160
GGGCCACCAG	CTCCAGACT	CGGCCATCAC	CCCGCAGAAA	GGAAAGGCCA	CAACCCGCGC	5220
ACCCAGCCCC	CGATCCGGCG	GGGAGCCACC	CAACCCGAAC	CAGCACCCAA	GAGCGATCCC	5280
CGAAGGACCC	CCGAACCGCA	AAGGACATCA	GTATCCCACA	GCCTCTCCAA	GTCCCCCGGT	5340
CTCTCCCCCT	TCTCGAAGGG	ACCAAAAGAT	CAATCCACCA	CACCCGACGA	CACTCAACTC	5400
CCCACCCCTA	AAGGAGACAC	CGGGAATCCC	AGAATCAAGA	CTCATCCAAT	GTCCATCATG	5460
GGTCTCAAGG	TGAACGTCTC	TGCCATATTC	ATGGCAGTAC	TGTTAACTCT	CCAAACACCC	5520
ACCGGTCAAA	TCCATTGGGG	CAATCTCTCT	AAGATAGGGG	TGGTAGGAAT	AGGAAGTGCA	5580
AGCTACAAAG	TTATGACTCG	TTCCAGCCAT	CAATCATTAG	TCATAAAATT	AATGCCCAAT	5640
ATAACTCTCC	TCAATAACTG	CACGAGGGTA	GAGATTGCAG	AATACAGGAG	ACTACTGAGA	5700
ACAGTTTTTG	AACCAATTAG	AGATGCACTT	AATGCAATGA	CCCAGAATAT	AAGACCGGTT	5760
CAGAGTGTAG	CTTCAAGTAG	GAGACACAAG	AGATTGCGG	GAGTAGTCCT	GGCAGGTGCG	5820
GCCCTAGGCG	TTGCCACAGC	TGCTCAGATA	ACAGCCGGCA	TTGCACTTCA	CCAGTCCATG	5880
CTGAACTCTC	AAGCCATCGA	CAATCTGAGA	GCGAGCCTGG	AACTACTAA	TCAGGCAATT	5940

SUBSTITUTE SHEET (RULE 26)

- 203 -

GAGGCAATCA GACAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC	6000
ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTAAAT CGGCCAGAAG	6060
CTCGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CATTATTGG CCCCAGCTTA	6120
CGGGACCCCA TATCTGCGGA GATATCTATC CAGGCTTGA GCTATGCGCT TGGAGGAGAC	6180
ATCAATAAGG TGTTAGAAAA GCTCGGATAC AGTGGAGGTG ATTTACTGGG CATCTTAGAG	6240
AGCAGAGGAA TAAAGGCCCG GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTCTC	6300
AGTATAGCCT ATCCGACGCT GTCCGAGATT AAGGGGGTGA TTGTCCACCG GCTAGAGGGG	6360
GTCTCGTACA ACATAGGCTC TCAAGAGTGG TATACCACTG TGCCCAAGTA TGTTGCAACC	6420
CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT	6480
GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TCCACCAAGT CCTGTGCTCG TACACTCGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TCACTGCCCG	6720
GTAGTCGAGG TGAACGGCGT GACCATCCAA GTCGGGAGCA GGAGGTATCC AGATGCTGTG	6780
TACTTGACCA GAATTGACCT CGGTCTCCC ATATCATTGG AGAGGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTAGCA TAGTCTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACGGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCACCCA GCATCAAGCC CACCTGAAAT	7200
TATCTCCGGC TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATTAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA	7320
TCCCAAGGGA AGTAGGATAG TCATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTGTTTG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500

SUBSTITUTE SHEET (RULE 26)

- 204 -

TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTCACTGACC TAGTGAAATT	7620
CATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCCTGTTAGA CTTGTATTTA GGTCGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTATGGGGGA ACTTACCTAG TGGAAAAGCC	7980
TAATCTGAGC AGCAAAAGGT CAGAGTTGTC ACAACTGAGC ATGTACCGAG TGTTTGAAGT	8040
AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATCTTGA	8100
GCAACCAGCC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAACTCGC	8160
AGCCCTTTGT CACGGGGAAG ATTCTATCAC AATTCCCTAT CAGGGATCAG GGAAAGGTGT	8220
CAGCTTCCAG CTCGTCAAGC TAGGTGTCCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CCCCTTATCA ACGGATGATC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGAACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAAC AGGCGTGTA GGGTAAAATC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTGATCT	8520
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA	8580
CGGTTCAAGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGAATATCCC	8640
GCCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTTAGTCCC TACCTCTTCA ATGTCCCAAT TAAGGAAGCA GGCAGAACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAAATC AGTTCCAATC TGGTGATTCT	8820
ACCTGGTCAA GATCTCCAAT ATGTTTTGGC AACCTACGAT ACTTCCAGGG TTGAACATGC	8880
TGTGGTTTAT TACGTTTACA GCCCAGGCCG CTCATTTTCT TACTTTTATC CTTTTAGGTT	8940
GCCTATAAAG GGGGTCCCCA TCGAATTACA AGTGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC	9060

SUBSTITUTE SHEET (RULE 26)

- 205 -

TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATCGCAG	9120
ATAGGGCTGC TAGTGAACCA ATCTCATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACCGCC TAAAAACGG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCACT	9480
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540
CGAGGAAGAT CCGTGAATC CTCAAAAAGG GGAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCTTAAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAACCTTG GAGTTTACAT GCACAGCTCC CAGTGGTTTG	9720
AGCCCTTTCT GTTTTGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACATTTGA ACTGGTTTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTCACTTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCATGAG TTAATTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA	10260
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT	10560
TTGCTGGAGT GAAATTTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT	10620

- 206 -

ACCTAAAGGA CAAGGCACTT GCTGCTCTCC AAAGGGAATG GGATTTCAGTT TACCCGAAAG 10680  
AGTTCCTGCG TTACGACCCCT CCCAAGGGAA CCGGGTCACG GAGGCTTGTA GATGTTTTCC 10740  
TTAATGATTG GAGCTTTGAC CCATATGATG TGATAATGTA TGTTGTAAGT GGAGCTTACC 10800  
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG 10860  
GTAGACTTTT TGCTAAAATG ACTTACAAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC 10920  
TAATCTCAAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCACG 10980  
ATTTGACTAA GGCACCTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAT CTCAAAGAAA 11040  
GTCACAGGGG GGGGCCAGTC TTA AAAACCT ACTCCCGAAG CCCAGTCCAC ACAAGTACCA 11100  
GGAACGTGAG AGCAGCAAAA GGGTTTATAG GGTTCCTCA AGTAATTCGG CAGGACCAAG 11160  
ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATTT ATCAGCACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTG CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCCTGTA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCAT ATCCCGTTAT 11400  
ATAAAGTCCC CAATGATCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATTG CCTATCTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCCTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CTTGTCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTCTGATCT 11940  
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCCTCACTAA TGCCTGAAGA GACCCCTCCAT CAAGTAATGA 12180

SUBSTITUTE SHEET (RULE 26)

- 207 -

CACAACAACC GGGGGACTCT TCATTCCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240  
TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA 12300  
TCCATAGTCC AAACCCAATG TTAAGAGGAT TATTCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC GGCATTCTCT ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCA CA GGGGCAAGAG AGTCTATTGC AGGCATGCTG GATACCACAA 12480  
AAGGCCTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGACTATGAA CAATTCAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GGCGAGAGCT CTAAGAAGCC 12660  
ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTTACGG CCTTGAGGTC CCTGATGTAC 12720  
TAGAATCTAT GCGAGGCCAC CTTATTCGGC GTCATGAGAC ATGTGTCATC TGCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTGTGTC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTCGTAAGA GCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA 12960  
CAGTGTACTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG 13020  
CTAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCGCATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTCGG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT 13200  
CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAAGG AATGCTTCTA GGGTTGGGTG 13260  
TTTTAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TGCCTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CCCACAAGCT AGAGCTGAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATACA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTGG 13500  
AATTGTGTTAC ATGGTCCACA CCCCAACTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CCTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAAT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GCTCATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740

SUBSTITUTE SHEET (RULE 26)

- 208 -

GACCATCAGG GAAATATCAG ATGGGTGAGC TGTTGTCATC GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGACACA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTCTTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA 14040  
GATTGACAAA CATCCAGGCA AAACACTTAT GTGTCTGGC AGATTGTAC TGTCAACCAG 14100  
GGACCTGCCC ACCAATTCGA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC 14160  
ATATCAAGGC AGAGGCTAGG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220  
TTGTAGACCA TTAATCATGC TCTCTGACTT ATCTCCGGCG AGGATCGATC AACAGATAA 14280  
GATTGAGAGT TGATCCAGGA TTCATTTTCG ACGCCCTCGC TGAGGTAAAT GTCAGTCAGC 14340  
CAAAGATCGG CAGCAACAAC ATCTCAAATA TGAGCATCAA GGCTTTCAGA CCCCCACACG 14400  
ATGATGTTGC AAAATTGCTC AAAGATATCA ACACAAGCAA GCACAACTTT CCCATTTTCAG 14460  
GGGGCAATCT CGCCAATTAT GAAATCCATG CTTTCCGCAG AATCGGGTTG AACTCATCTG 14520  
CTTGCTACAA AGCTGTTGAG ATATCAACAT TAATTAGGAG ATGCCTTGAG CCAGGGGAGG 14580  
ACGGCTTGTT CTTGGGTGAG GGATCGGGTT CTATGTTGAT CACTTATAAG GAGATACTTA 14640  
AACTAAACAA GTGCTTCTAT AATAGTGGGG TTTCCGCCAA TTCTAGATCT GGTCAAAGGG 14700  
AATTAGCACC CTATCCCTCC GAAGTTGGCC TTGTCGAACA CAGAATGGGA GTAGGTAATA 14760  
TTGTCAAAGT GCTCTTTAAC GGGAGGCCCG AAGTCACGTG GGTAGGCAGT GTAGATTGCT 14820  
TCAATTTTCAT AGTTAGTAAT ATCCCTACCT CTAGTGTGGG GTTTATCCAT TCAGATATAG 14880  
AGACCTTGCC TAACAAAGAT ACTATAGAGA AGCTAGAGGA ATTGGCAGCC ATCTTATCGA 14940  
TGGCTCTGCT CCTGGGCAAA ATAGGATCAA TACTGGTGAT TAAGCTTATG CCTTTCAGCG 15000  
GGGATTTTGT TCAGGGATTT ATAAGTTATG TAGGGTCCCA TTATAGAGAA GTGAACCTTG 15060  
TATACCCTAG ATACAGCAAC TTCATATCTA CTGAATCTTA TTTGGTTATG ACAGATCTCA 15120  
AGGCTAACCG GCTAATGAAT CCTGAAAAGA TTAAGCAGCA GATAATTGAA TCATCTGTGA 15180  
GGACTTCACC TGGACTTATA GGTCAATCC TATCCATTAA GCAACTAAGC TGCATACAAG 15240  
CAATTGTGGG AGACGCAGTT AGTAGAGGTG ATATCAATCC TACTCTGAAA AAACCTACAC 15300

SUBSTITUTE SHEET (RULE 26)



- 209 -

CTATAGAGCA GGTGCTGATC AATTGCGGGT TGGCAATTAA CGGACCTAAG CTGTGCAAAG 15360  
 AATTGATCCA CCATGATGTT GCCTCAGGGC AAGATGGATT GCTTAATTCT ATACTCATCC 15420  
 TCTACAGGGA GTTGGCAAGA TTCAAAGACA ACCAAAGAAG TCAACAAGGG ATGTTCCACG 15480  
 CTTACCCCGT ATTGGTAAGT AGCAGGCAAC GAGAACTTAT ATCTAGGATC ACCCGCAAAT 15540  
 TTTGGGGGCA CATTCTTCTT TACTCCGGGA ACAGAAAGTT GATAAATAAG TTTATCCAGA 15600  
 ATCTCAAGTC CGGCTATCTG ATACTAGACT TACACCAGAA TATCTTCGTT AAGAATCTAT 15660  
 CCAAGTCAGA GAAACAGATT ATTATGACGG GGGGTTTGAA ACGTGAGTGG GTTTTAAAGG 15720  
 TAACAGTCAA GGAGACCAAA GAATGGTATA AGTTAGTCGG ATACAGTGCC CTGATTAAGG 15780  
 ACTAATTGGT TGAAGTCCGG AACCTAATC CTGCCCTAGG TGGTTAGGCA TTATTTGCAA 15840  
 TAGATTAAAG AAAAATTGTA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT 15894

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
 1 5 10 15  
 Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
 20 25 30  
 Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
 35 40 45  
 Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
 50 55 60  
 Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
 65 70 75 80  
 Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn

SUBSTITUTE SHEET (RULE 26)

- 210 -

85	90	95
Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys		
100	105	110
Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu		
115	120	125
Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp		
130	135	140
Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln		
145	150	155
Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg		
165	170	175
Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr		
180	185	190
Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp		
195	200	205
Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr		
210	215	220
Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met		
225	230	235
Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly		
245	250	255
Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu		
260	265	270
Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu		
275	280	285
Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe		
290	295	300
Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly		
305	310	315
Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr		
325	330	335
Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe		
340	345	350
Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu		
355	360	365

SUBSTITUTE SHEET (RULE 26)

- 211 -

Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr  
 370 375 380  
 Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr  
 385 390 395 400  
 Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His  
 405 410 415  
 Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr  
 420 425 430  
 His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe  
 435 440 445  
 Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu  
 450 455 460  
 Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr  
 465 470 475 480  
 Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg  
 485 490 495  
 Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp  
 500 505 510  
 Val Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe  
 515 520 525  
 Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg  
 530 535 540  
 Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala  
 545 550 555 560  
 Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly  
 565 570 575  
 Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala  
 580 585 590  
 Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro  
 595 600 605  
 Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn  
 610 615 620  
 Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Val Ile Arg Gln  
 625 630 635 640

SUBSTITUTE SHEET (RULE 26)

- 212 -

Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val  
 645 650 655  
 Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670  
 Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly  
 675 680 685  
 Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val  
 690 695 700  
 Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile  
 705 710 715 720  
 Pro Leu Tyr Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met  
 725 730 735  
 Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile  
 740 745 750  
 Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser  
 755 760 765  
 Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro  
 770 775 780  
 Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr  
 785 790 795 800  
 Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His  
 805 810 815  
 His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr  
 820 825 830  
 Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845  
 Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860  
 Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880  
 Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895  
 Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910  
 Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile

SUBSTITUTE SHEET (RULE 26)

- 213 -

915	920	925
Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn 930	935	940
Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser 945	950	955 960
Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu 965	970	975
Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu 980	985	990
Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser 995	1000	1005
Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His 1010	1015	1020
Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu 1025	1030	1035 1040
Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val 1045	1050	1055
Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg 1060	1065	1070
Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala 1075	1080	1085
Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser 1090	1095	1100
Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly 1105	1110	1115 1120
Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu 1125	1130	1135
Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg 1140	1145	1150
Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly 1155	1160	1165
His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser 1170	1175	1180
Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp 1185	1190	1195 1200

SUBSTITUTE SHEET (RULE 26)

- 214 -

Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr  
 1205 1210 1215  
 Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser  
 1220 1225 1230  
 Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala  
 1235 1240 1245  
 Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg  
 1250 1255 1260  
 Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile  
 1265 1270 1275 1280  
 Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln  
 1285 1290 1295  
 Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr  
 1300 1305 1310  
 Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp  
 1315 1320 1325  
 Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu  
 1330 1335 1340  
 Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val  
 1345 1350 1355 1360  
 Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp  
 1365 1370 1375  
 His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu  
 1380 1385 1390  
 Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp  
 1395 1400 1405  
 Thr Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe  
 1410 1415 1420  
 Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr  
 1425 1430 1435 1440  
 Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met  
 1445 1450 1455  
 Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile  
 1460 1465 1470

SUBSTITUTE SHEET (RULE 26)

- 215 -

Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly  
 1475 1480 1485

Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro  
 1490 1495 1500

Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg  
 1505 1510 1515 1520

Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro  
 1525 1530 1535

Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His  
 1540 1545 1550

Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met  
 1555 1560 1565

Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu  
 1570 1575 1580

Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val  
 1585 1590 1595 1600

Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala  
 1605 1610 1615

Asp Leu Tyr Cys Gln Pro Gly Thr Cys Pro Pro Ile Arg Gly Leu Arg  
 1620 1625 1630

Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala  
 1635 1640 1645

Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val  
 1650 1655 1660

Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
 1665 1670 1675 1680

Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
 1685 1690 1695

Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
 1700 1705 1710

Met Ser Ile Lys Ala Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
 1715 1720 1725

Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
 1730 1735 1740

Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn

SUBSTITUTE SHEET (RULE 26)

- 216 -

1745	1750	1755	1760
Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg	1765	1770	1775
Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly	1780	1785	1790
Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe	1795	1800	1805
Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu	1810	1815	1820
Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val	1825	1830	1835
Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp	1845	1850	1855
Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr	1860	1865	1870
Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys	1875	1880	1885
Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala	1890	1895	1900
Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro	1905	1910	1915
Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser His	1925	1930	1935
Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser	1940	1945	1950
Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met	1955	1960	1965
Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr	1970	1975	1980
Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys	1985	1990	1995
Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro	2005	2010	2015
Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly	2020	2025	2030

SUBSTITUTE SHEET (RULE 26)



- 217 -

Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp  
 2035 2040 2045  
 Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr  
 2050 2055 2060  
 Arg Glu Leu Ala Arg Phe Lys Asp Asn Gln Arg Ser Gln Gln Gly Met  
 2065 2070 2075 2080  
 Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile  
 2085 2090 2095  
 Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly  
 2100 2105 2110  
 Asn Arg Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr  
 2115 2120 2125  
 Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
 2130 2135 2140  
 Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
 2145 2150 2155 2160  
 Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
 2165 2170 2175  
 Tyr Ser Ala Leu Ile Lys Asp  
 2180

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ACCAAACAAA GTTGGGTAAG GATAGTTCAA TCAATGATCA TCTTCTAGTG CACTTAGGAT	60
TCAAGATCCT ATTATCAGGG ACAAGAGCAG GATTAGGGAT ATCCGAGATG GCCACACTTT	120
TAAGGAGCTT AGCATTGTTC AAAAGAAACA AGGACAAACC ACCCATTACA TCAGGATCCG	180

SUBSTITUTE SHEET (RULE 26)

- 218 -

GTGGAGCCAT CAGAGGAATC AAACACATTA TTATAGTACC AATCCCTGGA GATTCTCTCAA	240
TTACCACTCG ATCCAGACTT CTGGACCGGT TGGTCAGGTT AATTGGAAAC CCGGATGTGA	300
GCGGGCCCCAA ACTAACAGGG GCACTAATAG GTATATTATC CTTATTTGTG GAGTCTCCAG	360
GTCAATTGAT TCAGAGGATC ACCGATGACC CTGACGTTAG CATAAGGCTG TTAGAGGTTG	420
TCCAGAGTGA CCAGTCACAA TCTGGCCTTA CCTTCGCATC AAGAGGTACC AACATGGAGG	480
ATGAGGCGGA CAAATACTTT TCACATGATG ATCCAATTAG TAGTGATCAA TCCAGGTTCTG	540
GATGGTTCTGA GAACAAGGAA ATCTCAGATA TTGAAGTGCA AGACCCTGAG GGATTCAACA	600
TGATTCTGGG TACCATCCTA GCCCAAATTT GGGTCTTGCT CGCAAAGGCG GTTACGGCCC	660
CAGACACGGC AGCTGATTCTG GAGCTAAGAA GGTGGATAAA GTACACCCAA CAAAGAAGGG	720
TAGTTGGTGA ATTTAGATTG GAGAGAAAAT GGTGGATGT GGTGAGGAAC AGGATTGCCG	780
AGGACCTCTC CTTACGCCGA TTCATGGTCG CTCTAATCCT GGATATCAAG AGAACACCCG	840
GAAACAAACC CAGGATTGCT GAAATGATAT GTGACATTGA TACATATATC GTAGAGGCAG	900
GATTAGCCAG TTTTATCCTG ACTATTAAGT TTGGGATAGA AACTATGTAT CCTGCTCTTG	960
GACTGCATGA ATTTGCTGGT GAGTTATCCA CACTTGAGTC CTTGATGAAC CTTTACCAGC	1020
AAATGGGGGA AACTGCACCC TACATGGTAA TCCTGGAGAA CTCAATTCAG AACAAGTTCA	1080
GTGCAGGATC ATACCCCTCTG CTCTGGAGCT ATGCCATGGG AGTAGGAGTG GAACTTGAAA	1140
ACTCCATGGG AGGTTTGAAC TTTGGCCGAT CTTACTTTGA TCCAGCATAT TTTAGATTAG	1200
GGCAAGAGAT GGTAAGGAGG TCAGCTGGAA AGGTCAGTTC CACATTGGCA TCTGAACTCG	1260
GTATCACTGC CGAGGATGCA AGGCTTGTTT CAGAGATTGC AATGCATACT ACTGAGGACA	1320
AGATCAGTAG AGCGGTTGGA CCCAGACAAG CCCAAGTATC ATTTCTACAC GGTGATCAAA	1380
GTGAGAATGA GCTACCGAGA TTGGGGGGCA AGGAAGATAG GAGGGTCAAA CAGAGTCGAG	1440
GAGAAGCCAG GGAGAGCTAC AGAGAAACCG GGCCCAGCAG AGCAAGTGAT GCGAGAGCTG	1500
CCCATCTTCC AACCGGCACA CCCCTAGACA TTGACACTGC ATCGGAGTCC AGCCAAGATC	1560
CGCAGGACAG TCGAAGGTCA GCTGACGCCC TGCTTAGGCT GCAAGCCATG GCAGGAATCT	1620
CGGAAGAACA AGGCTCAGAC ACGGACACCC CTATAGTGTA CAATGACAGA AATCTTCTAG	1680
ACTAGGTGCG AGAGGCCGAG GACCAGAACA ACATCCGCCT ACCCTCCATC ATTGTTATAA	1740

SUBSTITUTE SHEET (RULE 26)

- 219 -

AAAACCTTAGG AACCAGGTCC ACACAGCCGC CAGCCCATCA ACCATCCACT CCCACGATTG 1800  
GAGCCGATGG CAGAAGAGCA GGCACGCCAT GTCAAAAACG GACTGGAATG CATCCGGGCT 1860  
CTCAAGGCCG AGCCCATCGG CTCACTGGCC ATCGAGGAAG CTATGGCAGC ATGGTCAGAA 1920  
ATATCAGACA ACCCAGGACA GGAGCGAGCC ACCTGCAGGG AAGAGAAGGC AGGCAGTTCTG 1980  
GGTCTCAGCA AACCATGCCT CTCAGCAATT GGATCAACTG AAGGCGGTGC ACCTCGCATC 2040  
CGCGGTCAGG GACCTGGAGA GAGCGATGAC GACGCTGAAA CTTTGGGAAT CCCCCAAGA 2100  
AATCTCCAGG CATCAAGCAC TGGGTTACAG TGTTATTATG TTTATGATCA CAGCGGTGAA 2160  
GCGGTTAAGG GAATCCAAGA TGCTGACTCT ATCATGGTTC AATCAGGCCT TGATGGTGAT 2220  
AGCACCTTAT CAGGAGGAGA CAATGAATCT GAAAACAGCG ATGTGGATAT TGGCGAACCT 2280  
GATACCGAGG GATATGCTAT CACTGACCGG GGATCTGCTC CCATCTCTAT GGGGTTTCAGG 2340  
GCTTCTGATG TTGAAACTGC AGAAGGAGGG GAGATCCACG AGCTCCTGAG ACTCCAATCC 2400  
AGAGGCAACA ACTTTCCGAA GCTTGGGAAA ACTCTCAATG TTCCTCCGCC CCCGGACCCC 2460  
GGTAGGGCCA GCACTTCCGG GACACCCATT AAAAAGGGCA CAGACGCGAG ATTAGCCTCA 2520  
TTTGGAAACG AGATCGCGTC TTTATTGACA GGTGGTGCAA CCCAATGTGC TCGAAAGTCA 2580  
CCCTCGGAAC CATCAGGGCC AGGTGCACCT GCGGGGAATG TCCCCGAGTA TGTGAGCAAT 2640  
GCCGCACTGA TACAGGAGTG GACACCCGAA TCTGGTACCA CAATCTCCCC GAGATCCAG 2700  
AATAATGAAG AAGGGGGAGA CTATTATGAT GATGAGCTGT TCTCTGATGT CCAAGATATT 2760  
AAAACAGCCT TGGCCAAAAT ACACGAGGAT AATCAGAAGA TAATCTCCAA GCTAGAATCA 2820  
CTGCTGTTAT TGAAGGGAGA AGTTGAGTCA ATTAAGAAGC AGATCAACAG GCAAAATATC 2880  
AGCATATCCA CCCTGGAAGG ACACCTCTCA AGCATCATGA TCGCCATTCC TGGACTTGGG 2940  
AAGGATCCCA ACGACCCAC TGCAGATGTC GAAATCAATC CCGACTTGAA ACCCATCATA 3000  
GGCAGAGATT CAGGCCGAGC ACTGGCCGAA GTTCTCAAGA AACCCGTTGC CAGCCGACAA 3060  
CTCCAAGGAA TGACAAATGG ACGGACCAGT TCCAGAGGAC AGCTGCTGAA GGAATTTTCAG 3120  
CCAAAGCCGA TCGGGAAAAA GATGAGCTCA GCCGTCGGGT TTGTTCTGA CACCGGCCCT 3180  
GCATCAGCA GTGTAATCCG CTCCATTATA AAATCCAGCC GGCTAGAGGA GGATCGGAAG 3240  
CGTTACCTGA TGA CTCTCCT TGATGATATC AAAGGAGCCA ATGATCTTGC CAAGTTCCAC 3300

SUBSTITUTE SHEET (RULE 26)

- 220 -

CAGATGCTGA TGAAGATAAT AATGAAGTAG CTACAGCTCA ACTTACCTGC CAACCCCATG 3360  
CCAGTCGACC CAACTAGTAC AACCTAAATC CATTATAAAA AACTTAGGAG CAAAGTGATT 3420  
GCCTCCCAAG TTCCACAATG ACAGAGATCT ACGACTTCGA CAAGTCGGCA TGGGACATCA 3480  
AAGGGTCGAT CGCTCCGATA CAACCGACCA CCTACAGTGA TGGCAGGCTG GTGCCCCAGG 3540  
TCAGAGTCAT AGATCCTGGT CTAGGCGACA GGAAGGATGA ATGCTTTATG TACATGTCTC 3600  
TGCTGGGGGT TGTGAGGAC AGCGATCCCC TAGGGCCTCC AATCGGGCGA GCATTGTTGGT 3660  
CCCTGCCCTT AGGTGTTGGC AGATCCACAG CAAAGCCCCG AAAACTCCTC AAAGAGGCCA 3720  
CTGAGCTTGA CATAGTTGTT AGACGTACAG CAGGGCTCAA TGAAAACTG GTGTTCTACA 3780  
ACAACACCCC ACTAACTCTC CTCACACCTT GGAGAAAGGT CCTAACAACA GGGAGTGTCT 3840  
TCAACGCAAA CCAAGTGTGC AATGCGGTTA ATCTGATACC GCTCGATACC CCGCAGAGGT 3900  
TCCGTGTTGT TTATATGAGC ATCACCCGTC TTTCGGATAA CGGGTATTAC ACCGTTCCCTA 3960  
GAAGAATGCT GGAATTCAGA TCGGTCAATG CAGTGGCCTT CAACCTGCTG GTGACCCTTA 4020  
GGATTGACAA GGCATAGGC CCTGGGAAGA TCATCGACAA TACAGAGCAA CTTCTGAGG 4080  
CAACATTTAT GGTCCACATC GGGAACTTCA GGAGAAAGAA GAGTGAAGTC TACTCTGCCG 4140  
ATTATTGCAA AATGAAAATC GAAAAGATGG GCCTGGTTTT TGCATTGGT GGGATAGGGG 4200  
GCACCACTCT TCACATTAGA AGCACAGGCA AAATGAGCAA GACTCTCCAT GCACAACTCG 4260  
GGTTCAAGAA GACCTTATGT TACCCGCTGA TGGATATCAA TGAAGACCTT AATCGATTAC 4320  
TCTGGAGGAG CAGATGCAAG ATAGTAAGAA TCCAGGCAGT TTTGCAGCCA TCAGTTCCTC 4380  
AAGAATTCCG CATTTACGAC GACGTGATCA TAAATGATGA CCAAGGACTA TTCAAAGTTC 4440  
TGTAGACCGT AGTGCCCGAGC AATGCCCGAA AACGACCCCC CTCACAATGA CAGCCAGAAG 4500  
GCCCGGACAA AAAAGCCCCC TCCGAAAGAC TCCACTGACC AAGCGAGAGG CCAGCCAGCA 4560  
GCCGACGGCA AGCACGAACA CCAGGCGGCC CCAGCACAGA ACAGCCCTGA TACAAGGCCA 4620  
CCACCAGCCA CCCCATCTG CATCCTCCTC GTGGGACCCC CGAGGACCAA CCCCCAAGGC 4680  
TGCCCCCGAT CCAAACCACC AACC GCATCC CCACCACCCC CGGGAAAGAA ACCCCCAGCA 4740  
ATTGGAAGGC CCCTCCCCCT CTTCTCAAC ACAAGAACTC CACAACCGAA CCGCACAAGC 4800  
GACCGAGGTG ACCCAACCGC AGGCATCCGA CTCCTAGAC AGATCCTCTC TCCCCGGCAA 4860

SUBSTITUTE SHEET (RULE 26)

- 221 -

ACTAAACAAA ACTTAGGGCC AAGGAACATA CACACCCAAC AGAACCCAGA CCCC GGCCCA 4920  
CGGCGCCGCG CCCCCAACCC CCGACAACCA GAGGGAGCCC CCAACCAATC CCGCCGGCTC 4980  
CCCCGGTGCC CACAGGCAGG GACACCAACC CCCGAACAGA CCCAGCACCT AACCATCGAC 5040  
AATCCAAGAC GGGGGGGCCC CCCCCAAAAA AGGCCCCCAG GGGCCGACAG CCAGCACCGC 5100  
GAGGAAGCCC ACCCACCCCA CACACGACCA CGGCAACCAA ACCAGAACCC AGACCACCCT 5160  
GGGCCACCAG CTCCCAGACT CGGCCATCAC CCCGAGAAA GGAAAGGCCA CAACCCGCGC 5220  
ACCCAGCCC CGATCCGGCG GGGAGCCACC CAACCCGAAC CAGCACCCAA GAGCGATCCC 5280  
CGAAGGACCC CCGAACCGCA AAGGACATCA GTATCCACA GCCTCTCAA GTCCCCGGT 5340  
CTCCTCCTCT TCTCGAAGGG ACCAAAAGAT CAATCCACCA CACCCGACGA CACTCAACTC 5400  
CCCACCCCTA AAGGAGACAC CGGGAATCCC AGAATCAAGA CTCATCCAAT GTCCATCATG 5460  
GGTCTCAAGG TGAACGTCTC TGCCATATTC ATGGCAGTAC TGTAACTCT CAAACACCC 5520  
ACCGGTCAAA TCCATTGGGG CAATCTCTCT AAGATAGGGG TGGTAGGAAT AGGAAGTGCA 5580  
AGCTACAAAG TTATGACTCG TTCCAGCCAT CAATCATTAG TCATAAAATT AATGCCCAAT 5640  
ATAACTCTCC TCAATAACTG CACGAGGGTA GAGATTGCAG AATACAGGAG ACTACTGAGA 5700  
ACAGTTTGG AACCAATTAG AGATGCACTT AATGCAATGA CCCAGAATAT AAGACCGGTT 5760  
CAGAGTGTAG CTTCAAGTAG GAGACACAAG AGATTGCGG GAGTAGTCCT GGCAGGTGCG 5820  
GCCCTAGGCG TTGCCACAGC TGCTCAGATA ACAGCCGGCA TTGCACTTCA CCAGTCCATG 5880  
CTGAACTCTC AAGCCATCGA CAATCTGAGA GCGAGCCTGG AAATACTAA TCAGGCAATT 5940  
GAGGCAATCA GACAAGCAGG GCAGGAGATG ATATTGGCTG TTCAGGGTGT CCAAGACTAC 6000  
ATCAATAATG AGCTGATACC GTCTATGAAC CAACTATCTT GTGATTTAAT CGGCCAGAAG 6060  
CTCGGGCTCA AATTGCTCAG ATACTATACA GAAATCCTGT CATTATTGG CCCCAGCTTA 6120  
CGGGACCCCA TATCTGCGGA GATATCTATC CAGGCTTTGA GCTATGCGCT TGGAGGAGAC 6180  
ATCAATAAGG TGTTAGAAAA GCTCGGATAC AGTGGAGGTG ATTTACTGGG CATCTTAGAG 6240  
AGCAGAGGAA TAAAGGCCCC GATAACTCAC GTCGACACAG AGTCCTACTT CATTGTCCTC 6300  
AGTATAGCCT ATCCGACGCT GTCCGAGATT AAGGGGGTGA TTGTCCACCG GCTAGAGGGG 6360  
GTCTCGTACA ACATAGGCTC TCAAGAGTGG TATACCACTG TGCCCAAGTA TGTTGCAACC 6420

SUBSTITUTE SHEET (RULE 26)

- 222 -

CAAGGGTACC TTATCTCGAA TTTTGATGAG TCATCGTGTA CTTTCATGCC AGAGGGGACT	6480
GTGTGCAGCC AAAATGCCTT GTACCCGATG AGTCCTCTGC TCCAAGAATG CCTCCGGGGG	6540
TACACCAAGT CCTGTGCTCG TACACTCGTA TCCGGGTCTT TTGGGAACCG GTTCATTTTA	6600
TCACAAGGGA ACCTAATAGC CAATTGTGCA TCAATCCTTT GCAAGTGTTA CACAACAGGA	6660
ACGATCATT AATCAAGACCC TGACAAGATC CTAACATACA TTGCTGCCGA TAACTGCCCCG	6720
GTAGTCGAGG TGAACGGCGT GACCATCCAA GTCGGGAGCA GGAGGTATCC AGACGCTGTG	6780
TACTTGACCA GAATTGACCT CGGTCCTCCC ATATTATTGG AGAGGTTGGA CGTAGGGACA	6840
AATCTGGGGA ATGCAATTGC TAAGTTGGAG GATGCCAAGG AATTGTTGGA GTCATCGGAC	6900
CAGATATTGA GGAGTATGAA AGGTTTATCG AGCACTTGCA TAGTCTACAT CCTGATTGCA	6960
GTGTGTCTTG GAGGGTTGAT AGGGATCCCC GCTTTAATAT GTTGCTGCAG GGGGCGTTGT	7020
AACAAAAAGG GAGAACAAGT TGGTATGTCA AGACCAGGCC TAAAGCCTGA TCTTACGGGA	7080
ACATCAAAAT CCTATGTAAG GTCGCTCTGA TCCTCTACAA CTCTTGAAAC ACAAATGTCC	7140
CACAAGTCTC CTCTTCGTCA TCAAGCAACC ACCGCACCCA GCATCAAGCC CACCTGAAAT	7200
TATCTCCGGC TTCCCTCTGG CCGAACAATA TCGGTAGTTA ATTAAACTT AGGGTGCAAG	7260
ATCATCCACA ATGTCACCAC AACGAGACCG GATAAATGCC TTCTACAAAG ATAACCCCCA	7320
TCCCAAGGGA AGTAGGATAG TCATTAACAG AGAACATCTT ATGATTGATA GACCTTATGT	7380
TTTGCTGGCT GTTCTGTTTG TCATGTTTCT GAGCTTGATC GGGTTGCTAG CCATTGCAGG	7440
CATTAGACTT CATCGGGCAG CCATCTACAC CGCAGAGATC CATAAAAGCC TCAGCACCAA	7500
TCTAGATGTA ACTAACTCAA TCGAGCATCA GGTCAAGGAC GTGCTGACAC CACTCTTCAA	7560
AATCATCGGT GATGAAGTGG GCCTGAGGAC ACCTCAGAGA TTCACTGACC TAGTGAAATT	7620
CATCTCTGAC AAGATTAAAT TCCTTAATCC GGATAGGGAG TACGACTTCA GAGATCTCAC	7680
TTGGTGTATC AACCCGCCAG AGAGAATCAA ATTGGATTAT GATCAATACT GTGCAGATGT	7740
GGCTGCTGAA GAGCTCATGA ATGCATTGGT GAACTCAACT CTACTGGAGA CCAGAACAAC	7800
CAATCAGTTC CTAGCTGTCT CAAAGGGAAA CTGCTCAGGG CCCACTACAA TCAGAGGTCA	7860
ATTCTCAAAC ATGTCGCTGT CCTGTGTTAGA CTTGTATTTA GGTGAGGTT ACAATGTGTC	7920
ATCTATAGTC ACTATGACAT CCCAGGGAAT GTATGGGGGA ACTTACCTAG TGGAAAAGCC	7980

SUBSTITUTE SHEET (RULE 26)

- 223 -

TAATCTGAGC AGCAAAAGGT CAGAGTTGTC ACAAAGT GAGC ATGTACCGAG TGTGTTGAAGT	8040
AGGTGTTATC AGAAATCCGG GTTTGGGGGC TCCGGTGTTT CATATGACAA ACTATCTTGA	8100
GCAACCAGTC AGTAATGATC TCAGCAACTG TATGGTGGCT TTGGGGGAGC TCAAAGTTCG	8160
AGCCCTTTGT CACCGGGAAG ATTCTATCAC AATTCCTAT CAGGGATCAG GGAAAGGTGT	8220
CAGCTTCCAG CTCGTCAAGC TAGGTGTCTG GAAATCCCCA ACCGACATGC AATCCTGGGT	8280
CACCTTATCA ACGGATGATC CAGTGATAGA CAGGCTTTAC CTCTCATCTC ACAGAGGTGT	8340
TATCGCTGAC AATCAAGCAA AATGGGCTGT CCCGACAACA CGAACAGATG ACAAGTTGCG	8400
AATGGAGACA TGCTTCCAAC AGGCGTGTA GGGTAAATC CAAGCACTCT GCGAGAATCC	8460
CGAGTGGGCA CCATTGAAGG ATAACAGGAT TCCTTCATAC GGGGTCTTGT CTGTTGATCT	8520
GAGTCTGACA GTTGAGCTTA AAATCAAAAT TGCTTCGGGA TTCGGGCCAT TGATCACACA	8580
CGGTTCAAGG ATGGACCTAT ACAAATCCAA CCACAACAAT GTGTATTGGC TGAATATCCC	8640
ACCAATGAAG AACCTAGCCT TAGGTGTAAT CAACACATTG GAGTGGATAC CGAGATTCAA	8700
GGTTAGTCCC TACCTCTTCA ATGTCCCAAT TAAGGAAGCA GGCGAAGACT GCCATGCCCC	8760
AACATACCTA CCTGCGGAGG TGGATGGTGA TGTCAACTC AGTTCCAATC TGGTGATTCT	8820
ACCTGGTCAA GATCTCCAAT ATGTTTGGC AACCTACGAT ACTCCAGGG TTGAACATGC	8880
TGTGGTTTAT TACGTTTACA GCCCAAGCCG CTCATTTTCT TACTTTTATC CTTTATAGTT	8940
GCCTATAAAG GGGGTCCCCA TCGAATTACA AGTGGAATGC TTCACATGGG ACCAAAAACT	9000
CTGGTGCCGT CACTTCTGTG TGCTTGCGGA CTCAGAATCT GGTGGACATA TCACTCACTC	9060
TGGGATGGTG GGCATGGGAG TCAGCTGCAC AGTCACCCGG GAAGATGGAA CCAATCGCAG	9120
ATAGGGCTGC TAGTGAACCTA ATCTCATGAT GTCACCCAGA CATCAGGCAT ACCCACTAGT	9180
GTGAAATAGA CATCAGAATT AAGAAAAACG TAGGGTCCAA GTGGTTCCCC GTTATGGACT	9240
CGCTATCTGT CAACCAGATC TTATACCCTG AAGTTCACCT AGATAGCCCG ATAGTTACCA	9300
ATAAGATAGT AGCCATCCTG GAGTATGCTC GAGTCCCTCA CGCTTACAGC CTGGAGGACC	9360
CTACACTGTG TCAGAACATC AAGCACCACC TAAAAACCG ATTTTCCAAC CAAATGATTA	9420
TAAACAATGT GGAAGTTGGG AATGTCATCA AGTCCAAGCT TAGGAGTTAT CCGGCCCCACT	9480
CTCATATTCC ATATCCAAAT TGTAATCAGG ATTTATTTAA CATAGAAGAC AAAGAGTCAA	9540

SUBSTITUTE SHEET (RULE 26)

- 224 -

CGAGGAAGAT CCGTGAACTC CTCAAAAAGG GGAATTCGCT GTACTCCAAA GTCAGTGATA	9600
AGGTTTTCCA ATGCTTAAGG GACACTAACT CACGGCTTGG CCTAGGCTCC GAATTGAGGG	9660
AGGACATCAA GGAGAAAGTT ATTAAGTTGG GAGTTACAT GCACAGCTCC CAGTGGTTTG	9720
AGCCCTTTCT GTTTTGGTTT ACAGTCAAGA CTGAGATGAG GTCAGTGATT AAATCACAAA	9780
CCCATACTTG CCATAGGAGG AGACACACAC CTGTATTCTT CACTGGTAGT TCAGTTGAGT	9840
TGCTAATCTC TCGTGACCTT GTTGCTATAA TCAGTAAAGA GTCTCAACAT GTATATTACC	9900
TGACATTTGA ACTGGTTTTG ATGTATTGTG ATGTCATAGA GGGGAGGTTA ATGACAGAGA	9960
CCGCTATGAC TATTGATGCT AGGTATACAG AGCTTCTAGG AAGAGTCAGA TACATGTGGA	10020
AACTGATAGA TGGTTTCTTC CCTGCACTCG GGAATCCAAC TTATCAAATT GTAGCCATGC	10080
TGGAGCCTCT TTCACTTGCT TACCTGCAGC TGAGGGATAT AACAGTAGAA CTCAGAGGTG	10140
CTTTCCTTAA CCACTGCTTT ACTGAAATAC ATGATGTTCT TGACCAAAC GGGTTTTCTG	10200
ATGAAGGTAC TTATCATGAG TTAATTGAAG CTCTAGATTA CATTTTCATA ACTGATGACA	10260
TACATCTGAC AGGGGAGATT TTCTCATTTT TCAGAAGTTT CGGCCACCCC AGACTTGAAG	10320
CAGTAACGGC TGCTGAAAAT GTTAGGAAAT ACATGAATCA GCCTAAAGTC ATTGTGTATG	10380
AGACTCTGAT GAAAGGTCAT GCCATATTTT GTGGAATCAT AATCAACGGC TATCGTGACA	10440
GGCACGGAGG CAGTTGGCCA CCGCTGACCC TCCCCCTGCA TGCTGCAGAC ACAATCCGGA	10500
ATGCTCAAGC TTCAGGTGAA GGGTTAACAC ATGAGCAGTG CGTTGATAAC TGGAAATCTT	10560
TTGCTGGAGT GAAATTTGGC TGCTTTATGC CTCTTAGCCT GGATAGTGAT CTGACAATGT	10620
ACCTAAAGGA CAAGGCACCT GCTGCTCTCC AAAGGGAATG GGATTCAGTT TACCCGAAAG	10680
AGTTCCTGCG TTACGACCCT CCAAGGGAA CCGGGTCACG GAGGCTTGTA GATGTTTTCC	10740
TTAATGATTC GAGCTTTGAC CCATATGATG TGATAATGTA TGTGTAAGT GGAGCTTACC	10800
TCCATGACCC TGAGTTCAAC CTGTCTTACA GCCTGAAAGA AAAGGAGATC AAGGAAACAG	10860
GTAGACTTTT TGCTAAATG ACTTACAAA TGAGGGCATG CCAAGTGATT GCTGAAAATC	10920
TAATCTCAA CGGGATTGGC AAATATTTTA AGGACAATGG GATGGCCAAG GATGAGCACC	10980
ATTGACTAA GGCACCTCCAC ACTCTAGCTG TCTCAGGAGT CCCCAGAGAT CTCAAAGAAA	11040
GTCACAGGGG GGGGCCAGTC TTAAAAACCT ACTCCCGAAG CCCAGTCCAC ACAAGTACCA	11100

SUBSTITUTE SHEET (RULE 26)



- 225 -

GGAACGTGAG AGCAGCAAAA GGGTTTATAG GGTTCCTCA AGTAATTCGG CAGGACCAAG 11160  
ACACTGATCA TCCGGAGAAT ATGGAAGCTT ACGAGACAGT CAGTGCATTT ATCAGGACTG 11220  
ATCTCAAGAA GTACTGCCTT AATTGGAGAT ATGAGACCAT CAGCTTGTTT GCACAGAGGC 11280  
TAAATGAGAT TTACGGATTG CCCTCATTTT TCCAGTGGCT GCATAAGAGG CTTGAGACCT 11340  
CTGTCCTGTA TGTAAGTGAC CCTCATTGCC CCCCCGACCT TGACGCCCCAT ATCCCGTTAT 11400  
ATAAAGTCCC CAATGATCAA ATCTTCATTA AGTACCCTAT GGGAGGTATA GAAGGGTATT 11460  
GTCAGAAGCT GTGGACCATC AGCACCATTG CCTATCTATA CCTGGCTGCT TATGAGAGCG 11520  
GAGTAAGGAT TGCTTCGTTA GTGCAAGGGG ACAATCAGAC CATAGCCGTA AAAAAAGGG 11580  
TACCCAGCAC ATGGCCCTAC AACCTTAAGA AACGGGAAGC TGCTAGAGTA ACTAGAGATT 11640  
ACTTTGTAAT TCTTAGGCAA AGGCTACATG ATATTGGCCA TCACCTCAAG GCAAATGAGA 11700  
CAATTGTTTC ATCACATTTT TTTGTCTATT CAAAAGGAAT ATATTATGAT GGGCTACTTG 11760  
TGTCCCAATC ACTCAAGAGC ATCGCAAGAT GTGTATTCTG GTCAGAGACT ATAGTTGATG 11820  
AAACAAGGGC AGCATGCAGT AATATTGCTA CAACAATGGC TAAAAGCATC GAGAGAGGTT 11880  
ATGACCGTTA CTTTGCATAT TCCCTGAACG TCCTAAAAGT GATACAGCAA ATTCTGATCT 11940  
CTCTTGGCTT CACAATCAAT TCAACCATGA CCCGGGATGT AGTCATACCC CTCCTCACAA 12000  
ACAACGACCT CTTAATAAGG ATGGCACTGT TGCCCGCTCC TATTGGGGGG ATGAATTATC 12060  
TGAATATGAG CAGGCTGTTT GTCAGAAACA TCGGTGATCC AGTAACATCA TCAATTGCTG 12120  
ATCTCAAGAG AATGATTCTC GCCTCACTAA TGCCTGAAGA GACCCTCCAT CAAGTAATGA 12180  
CACAACAACC GGGGGACTCT TCATTCTAG ACTGGGCTAG CGACCCTTAC TCAGCAAATC 12240  
TTGTATGTGT CCAGAGCATC ACTAGACTCC TCAAGAACAT AACTGCAAGG TTTGTCCTGA 12300  
TCCATAGTCC AAACCCAATG TTAAAAGGAT TATTCCATGA TGACAGTAAA GAAGAGGACG 12360  
AGGGACTGGC GGCATTCTCT ATGGACAGGC ATATTATAGT ACCTAGGGCA GCTCATGAAA 12420  
TCCTGGATCA TAGTGTCA CAAGGCAAGAG AGTCTATTGC AGGCATGCTG GATACCACAA 12480  
AAGGCCTGAT TCGAGCCAGC ATGAGGAAGG GGGGGTTAAC CTCTCGAGTG ATAACCAGAT 12540  
TGTCCAATTA TGAATATGAA CAATTCAGAG CAGGGATGGT GCTATTGACA GGAAGAAAGA 12600  
GAAATGTCCT CATTGACAAA GAGTCATGTT CAGTGCAGCT GGCGAGAGCT CTAAGAAGCC 12660

SUBSTITUTE SHEET (RULE 26)

- 226 -

ATATGTGGGC GAGGCTAGCT CGAGGACGGC CTATTACGG CCTTGAGGTC CCTGATGTAC 12720  
TAGAATCTAT GCGAGGCCAC CTTATTCGGC GTCATGAGAC ATGTGTCATC TCGAGTGTG 12780  
GATCAGTCAA CTACGGATGG TTTTTGTCC CCTCGGGTTG CCAACTGGAT GATATTGACA 12840  
AGGAAACATC ATCCTTGAGA GTCCCATATA TTGGTTCTAC CACTGATGAG AGAACAGACA 12900  
TGAAGCTTGC CTTGTAAGA GCCCCAAGTC GATCCTTGCG ATCTGCTGTT AGAATAGCAA 12960  
CAGTGTACTC ATGGGCTTAC GGTGATGATG ATAGCTCTTG GAACGAAGCC TGGTTGTTGG 13020  
CTAGGCAAAG GGCCAATGTG AGCCTGGAGG AGCTAAGGGT GATCACTCCC ATCTCAACTT 13080  
CGACTAATTT AGCGCATAGG TTGAGGGATC GTAGCACTCA AGTGAAATAC TCAGGTACAT 13140  
CCCTTGTCGG AGTGGCGAGG TATACCACAA TCTCCAACGA CAATCTCTCA TTTGTCATAT 13200  
CAGATAAGAA GGTGATACT AACTTTATAT ACCAACAAGG AATGCTTCTA GGGTTGGGTG 13260  
TTTTAGAAAC ATTGTTTCGA CTCGAGAAAG ATACCGGATC ATCTAACACG GTATTACATC 13320  
TTCACGTCGA AACAGATTGT TGCCTGATCC CGATGATAGA TCATCCCAGG ATACCCAGCT 13380  
CCCGCAAGCT AGAGCTGAGG GCAGAGCTAT GTACCAACCC ATTGATATAT GATAATGCAC 13440  
CTTTAATTGA CAGAGATACA ACAAGGCTAT ACACCCAGAG CCATAGGAGG CACCTTGTTG 13500  
AATTTGTTAC ATGGTCCACA CCCCACCTAT ATCACATTTT AGCTAAGTCC ACAGCACTAT 13560  
CTATGATTGA CTTGGTAACA AAATTTGAGA AGGACCATAT GAATGAAATT TCAGCTCTCA 13620  
TAGGGGATGA CGATATCAAT AGTTTCATAA CTGAGTTTCT GTCATAGAG CCAAGATTAT 13680  
TCACTATCTA CTTGGGCCAG TGTGCGGCCA TCAATTGGGC ATTTGATGTA CATTATCATA 13740  
GACCATCAGG GAAATATCAG ATGGGTGAGC TGTGTCATC GTTCCTTTCT AGAATGAGCA 13800  
AAGGAGTGTT TAAGGTGCTT GTCAATGCTC TAAGCCACCC AAAGATCTAC AAGAAATTCT 13860  
GGCATTGTGG TATTATAGAG CCTATCCATG GTCCTTCACT TGATGCTCAA AACTTGACAA 13920  
CAACTGTGTG CAACATGGTT TACACATGCT ATATGACCTA CCTCGACCTG TTGTTGAATG 13980  
AAGAGTTAGA AGAGTTCACA TTTCTCTTGT GTGAAAGCGA CGAGGATGTA GTACCGGACA 14040  
GATTCGACAA CATCCAGGCA AAACACTTAT GTGTTCTGGC AGATTTGTAC TGTCAACCAG 14100  
GGGCTGCCC ACCAATTCGA GGTCTAAGAC CGGTAGAGAA ATGTGCAGTT CTAACCGACC 14160  
ATATCAAGGC AGAGGCTAGG TTATCTCCAG CAGGATCTTC GTGGAACATA AATCCAATTA 14220

SUBSTITUTE SHEET (RULE 26)

- 227 -

TTGTAGACCA	TTACTCATGC	TCTCTGACTT	ATCTCCGGCG	AGGATCGATC	AAACAGATAA	14280
GATTGAGAGT	TGATCCAGGA	TTCATTTTCG	ACGCCCTCGC	TGAGGTAAAT	GTCAGTCAGC	14340
CAAAGATCGG	CAGCAACAAC	ATCTCAAATA	TGAGCATCAA	GGCTTTCAGA	CCCCCACACG	14400
ATGATGTTGC	AAAATTGCTC	AAAGATATCA	ACACAAGCAA	GCACAATCTT	CCCATTTCAG	14460
GGGGCAATCT	CGCCAATTAT	GAAATCCATG	CTTTCGCAG	AATCGGGTTG	AACTCATCTG	14520
CTTGCTACAA	AGCTGTTGAG	ATATCAACAT	TAATTAGGAG	ATGCCTTGAG	CCAGGGGAGG	14580
ACGGCTTGTT	CTTGGGTGAG	GGATCGGGTT	CTATGTTGAT	CACTTATAAG	GAGATACTTA	14640
AACTAAACAA	GTGCTTCTAT	AATAGTGGGG	TTTCCGCCAA	TTCTAGATCT	GGTCAAAGGG	14700
AATTAGCACC	CTATCCCTCC	GAAATTGGCC	TTGTCGAACA	CAGAATGGGA	GTAGGTAATA	14760
TTGTCAAAGT	GCTCTTTAAC	GGGAGGCCCG	AAGTCACGTG	GGTAGGCAGT	GTAGATTGCT	14820
TCAATTTTCT	AGTTAGTAAT	ATCCCTACCT	CTAGTGTGGG	GTTTATCCAT	TCAGATATAG	14880
AGACCTTGCC	TAACAAAGAT	ACTATAGAGA	AGCTAGAGGA	ATTGGCAGCC	ATCTTATCGA	14940
TGGCTCTGCT	CCTGGGCAAA	ATAGGATCAA	TACTGGTGAT	TAAGCTTATG	CCTTTCAGCG	15000
GGGATTTTGT	TCAGGGATTT	ATAAGTTATG	TAGGGTCTTA	TTATAGAGAA	GTGAACCTTG	15060
TATACCCTAG	ATACAGCAAC	TTCATATCTA	CTGAATCTTA	TTTGGTTATG	ACAGATCTCA	15120
AGGCTAACCG	GCTAATGAAT	CCTGAAAAGA	TTAAGCAGCA	GATAATTGAA	TCATCTGTGA	15180
GGACTTCACC	TGGACTTATA	GGTCACATCC	TATCCATTAA	GCAACTAAGC	TGCATACAAG	15240
CAATTGTGGG	AGACGCAGTT	AGTAGAGGTG	ATATCAATCC	TACTCTGAAA	AAACTTACAC	15300
CTATAGAGCA	GGTGCTGATC	AATTGCGGGT	TGGCAATTAA	CGGACCTAAG	CTGTGCAAAG	15360
AATTGATCCA	CCATGATGTT	GCCTCAGGGC	AAGATGGATT	GCTTAATTCT	ATACTCATCC	15420
TCTACAGGGA	GTTGGCAAGA	TTCAAAGACA	ACCGAAGAAG	TCAACAAGGG	ATGTTCCACG	15480
CTTACCCCGT	ATTGGTAAGT	AGCAGGCAAC	GAGAACTTAT	ATCTAGGATC	ACCCGCAAAT	15540
TTTGGGGGCA	CATTCTTCTT	TACTCCGGGA	ACAGAAAGTT	GATAAATAAG	TTTATCCAGA	15600
ATCTCAAGTC	CGGCTATCTG	ATACTAGACT	TACACCAGAA	TATCTTCGTT	AAGAATCTAT	15660
CCAAGTCAGA	GAAACAGATT	ATTATGACGG	GGGGTTTGAA	ACGTGAGTGG	GTTTTTAAGG	15720
TAACAGTCAA	GGAGACCAAA	GAATGGTATA	AGTTAGTCGG	ATACAGTGCC	CTGATTAAAG	15780

SUBSTITUTE SHEET (RULE 26)

- 228 -

ACTAATTGAT TGAACCTCCGG AACCCCTAATC CTGCCCTAGG TGGTTAGGCA TTATTTGCAA 15840

TATATTAAAG AAAACTTTGA AAATACGAAG TTTCTATTCC CAGCTTTGTC TGGT 15894

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2183 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Asp Ser Leu Ser Val Asn Gln Ile Leu Tyr Pro Glu Val His Leu  
1 5 10 15

Asp Ser Pro Ile Val Thr Asn Lys Ile Val Ala Ile Leu Glu Tyr Ala  
20 25 30

Arg Val Pro His Ala Tyr Ser Leu Glu Asp Pro Thr Leu Cys Gln Asn  
35 40 45

Ile Lys His Arg Leu Lys Asn Gly Phe Ser Asn Gln Met Ile Ile Asn  
50 55 60

Asn Val Glu Val Gly Asn Val Ile Lys Ser Lys Leu Arg Ser Tyr Pro  
65 70 75 80

Ala His Ser His Ile Pro Tyr Pro Asn Cys Asn Gln Asp Leu Phe Asn  
85 90 95

Ile Glu Asp Lys Glu Ser Thr Arg Lys Ile Arg Glu Leu Leu Lys Lys  
100 105 110

Gly Asn Ser Leu Tyr Ser Lys Val Ser Asp Lys Val Phe Gln Cys Leu  
115 120 125

Arg Asp Thr Asn Ser Arg Leu Gly Leu Gly Ser Glu Leu Arg Glu Asp  
130 135 140

Ile Lys Glu Lys Val Ile Asn Leu Gly Val Tyr Met His Ser Ser Gln  
145 150 155 160

Trp Phe Glu Pro Phe Leu Phe Trp Phe Thr Val Lys Thr Glu Met Arg  
165 170 175

SUBSTITUTE SHEET (RULE 26)

- 229 -

Ser Val Ile Lys Ser Gln Thr His Thr Cys His Arg Arg Arg His Thr  
 180 185 190

Pro Val Phe Phe Thr Gly Ser Ser Val Glu Leu Leu Ile Ser Arg Asp  
 195 200 205

Leu Val Ala Ile Ile Ser Lys Glu Ser Gln His Val Tyr Tyr Leu Thr  
 210 215 220

Phe Glu Leu Val Leu Met Tyr Cys Asp Val Ile Glu Gly Arg Leu Met  
 225 230 235 240

Thr Glu Thr Ala Met Thr Ile Asp Ala Arg Tyr Thr Glu Leu Leu Gly  
 245 250 255

Arg Val Arg Tyr Met Trp Lys Leu Ile Asp Gly Phe Phe Pro Ala Leu  
 260 265 270

Gly Asn Pro Thr Tyr Gln Ile Val Ala Met Leu Glu Pro Leu Ser Leu  
 275 280 285

Ala Tyr Leu Gln Leu Arg Asp Ile Thr Val Glu Leu Arg Gly Ala Phe  
 290 295 300

Leu Asn His Cys Phe Thr Glu Ile His Asp Val Leu Asp Gln Asn Gly  
 305 310 315 320

Phe Ser Asp Glu Gly Thr Tyr His Glu Leu Ile Glu Ala Leu Asp Tyr  
 325 330 335

Ile Phe Ile Thr Asp Asp Ile His Leu Thr Gly Glu Ile Phe Ser Phe  
 340 345 350

Phe Arg Ser Phe Gly His Pro Arg Leu Glu Ala Val Thr Ala Ala Glu  
 355 360 365

Asn Val Arg Lys Tyr Met Asn Gln Pro Lys Val Ile Val Tyr Glu Thr  
 370 375 380

Leu Met Lys Gly His Ala Ile Phe Cys Gly Ile Ile Ile Asn Gly Tyr  
 385 390 395 400

Arg Asp Arg His Gly Gly Ser Trp Pro Pro Leu Thr Leu Pro Leu His  
 405 410 415

Ala Ala Asp Thr Ile Arg Asn Ala Gln Ala Ser Gly Glu Gly Leu Thr  
 420 425 430

His Glu Gln Cys Val Asp Asn Trp Lys Ser Phe Ala Gly Val Lys Phe  
 435 440 445

Gly Cys Phe Met Pro Leu Ser Leu Asp Ser Asp Leu Thr Met Tyr Leu

SUBSTITUTE SHEET (RULE 26)

- 230 -

450	455	460
Lys Asp Lys Ala Leu Ala Ala Leu Gln Arg Glu Trp Asp Ser Val Tyr		
465	470	475 480
Pro Lys Glu Phe Leu Arg Tyr Asp Pro Pro Lys Gly Thr Gly Ser Arg		
	485	490 495
Arg Leu Val Asp Val Phe Leu Asn Asp Ser Ser Phe Asp Pro Tyr Asp		
	500	505 510
Val Ile Met Tyr Val Val Ser Gly Ala Tyr Leu His Asp Pro Glu Phe		
	515	520 525
Asn Leu Ser Tyr Ser Leu Lys Glu Lys Glu Ile Lys Glu Thr Gly Arg		
	530	535 540
Leu Phe Ala Lys Met Thr Tyr Lys Met Arg Ala Cys Gln Val Ile Ala		
	545	550 555 560
Glu Asn Leu Ile Ser Asn Gly Ile Gly Lys Tyr Phe Lys Asp Asn Gly		
	565	570 575
Met Ala Lys Asp Glu His Asp Leu Thr Lys Ala Leu His Thr Leu Ala		
	580	585 590
Val Ser Gly Val Pro Lys Asp Leu Lys Glu Ser His Arg Gly Gly Pro		
	595	600 605
Val Leu Lys Thr Tyr Ser Arg Ser Pro Val His Thr Ser Thr Arg Asn		
	610	615 620
Val Arg Ala Ala Lys Gly Phe Ile Gly Phe Pro Gln Val Ile Arg Gln		
	625	630 635 640
Asp Gln Asp Thr Asp His Pro Glu Asn Met Glu Ala Tyr Glu Thr Val		
	645	650 655
Ser Ala Phe Ile Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg		
	660	665 670
Tyr Glu Thr Ile Ser Leu Phe Ala Gln Arg Leu Asn Glu Ile Tyr Gly		
	675	680 685
Leu Pro Ser Phe Phe Gln Trp Leu His Lys Arg Leu Glu Thr Ser Val		
	690	695 700
Leu Tyr Val Ser Asp Pro His Cys Pro Pro Asp Leu Asp Ala His Ile		
	705	710 715 720
Pro Leu Tyr Lys Val Pro Asn Asp Gln Ile Phe Ile Lys Tyr Pro Met		
	725	730 735

SUBSTITUTE SHEET (RULE 26)

- 231 -

Gly Gly Ile Glu Gly Tyr Cys Gln Lys Leu Trp Thr Ile Ser Thr Ile  
 740 745 750  
 Pro Tyr Leu Tyr Leu Ala Ala Tyr Glu Ser Gly Val Arg Ile Ala Ser  
 755 760 765  
 Leu Val Gln Gly Asp Asn Gln Thr Ile Ala Val Thr Lys Arg Val Pro  
 770 775 780  
 Ser Thr Trp Pro Tyr Asn Leu Lys Lys Arg Glu Ala Ala Arg Val Thr  
 785 790 795 800  
 Arg Asp Tyr Phe Val Ile Leu Arg Gln Arg Leu His Asp Ile Gly His  
 805 810 815  
 His Leu Lys Ala Asn Glu Thr Ile Val Ser Ser His Phe Phe Val Tyr  
 820 825 830  
 Ser Lys Gly Ile Tyr Tyr Asp Gly Leu Leu Val Ser Gln Ser Leu Lys  
 835 840 845  
 Ser Ile Ala Arg Cys Val Phe Trp Ser Glu Thr Ile Val Asp Glu Thr  
 850 855 860  
 Arg Ala Ala Cys Ser Asn Ile Ala Thr Thr Met Ala Lys Ser Ile Glu  
 865 870 875 880  
 Arg Gly Tyr Asp Arg Tyr Leu Ala Tyr Ser Leu Asn Val Leu Lys Val  
 885 890 895  
 Ile Gln Gln Ile Leu Ile Ser Leu Gly Phe Thr Ile Asn Ser Thr Met  
 900 905 910  
 Thr Arg Asp Val Val Ile Pro Leu Leu Thr Asn Asn Asp Leu Leu Ile  
 915 920 925  
 Arg Met Ala Leu Leu Pro Ala Pro Ile Gly Gly Met Asn Tyr Leu Asn  
 930 935 940  
 Met Ser Arg Leu Phe Val Arg Asn Ile Gly Asp Pro Val Thr Ser Ser  
 945 950 955 960  
 Ile Ala Asp Leu Lys Arg Met Ile Leu Ala Ser Leu Met Pro Glu Glu  
 965 970 975  
 Thr Leu His Gln Val Met Thr Gln Gln Pro Gly Asp Ser Ser Phe Leu  
 980 985 990  
 Asp Trp Ala Ser Asp Pro Tyr Ser Ala Asn Leu Val Cys Val Gln Ser  
 995 1000 1005

SUBSTITUTE SHEET (RULE 26)

- 232 -

Ile Thr Arg Leu Leu Lys Asn Ile Thr Ala Arg Phe Val Leu Ile His  
1010 1015 1020

Ser Pro Asn Pro Met Leu Lys Gly Leu Phe His Asp Asp Ser Lys Glu  
1025 1030 1035 1040

Glu Asp Glu Gly Leu Ala Ala Phe Leu Met Asp Arg His Ile Ile Val  
1045 1050 1055

Pro Arg Ala Ala His Glu Ile Leu Asp His Ser Val Thr Gly Ala Arg  
1060 1065 1070

Glu Ser Ile Ala Gly Met Leu Asp Thr Thr Lys Gly Leu Ile Arg Ala  
1075 1080 1085

Ser Met Arg Lys Gly Gly Leu Thr Ser Arg Val Ile Thr Arg Leu Ser  
1090 1095 1100

Asn Tyr Asp Tyr Glu Gln Phe Arg Ala Gly Met Val Leu Leu Thr Gly  
1105 1110 1115 1120

Arg Lys Arg Asn Val Leu Ile Asp Lys Glu Ser Cys Ser Val Gln Leu  
1125 1130 1135

Ala Arg Ala Leu Arg Ser His Met Trp Ala Arg Leu Ala Arg Gly Arg  
1140 1145 1150

Pro Ile Tyr Gly Leu Glu Val Pro Asp Val Leu Glu Ser Met Arg Gly  
1155 1160 1165

His Leu Ile Arg Arg His Glu Thr Cys Val Ile Cys Glu Cys Gly Ser  
1170 1175 1180

Val Asn Tyr Gly Trp Phe Phe Val Pro Ser Gly Cys Gln Leu Asp Asp  
1185 1190 1195 1200

Ile Asp Lys Glu Thr Ser Ser Leu Arg Val Pro Tyr Ile Gly Ser Thr  
1205 1210 1215

Thr Asp Glu Arg Thr Asp Met Lys Leu Ala Phe Val Arg Ala Pro Ser  
1220 1225 1230

Arg Ser Leu Arg Ser Ala Val Arg Ile Ala Thr Val Tyr Ser Trp Ala  
1235 1240 1245

Tyr Gly Asp Asp Asp Ser Ser Trp Asn Glu Ala Trp Leu Leu Ala Arg  
1250 1255 1260

Gln Arg Ala Asn Val Ser Leu Glu Glu Leu Arg Val Ile Thr Pro Ile  
1265 1270 1275 1280

Ser Thr Ser Thr Asn Leu Ala His Arg Leu Arg Asp Arg Ser Thr Gln

SUBSTITUTE SHEET (RULE 26)



- 233 -

1285	1290	1295
Val Lys Tyr Ser Gly Thr Ser Leu Val Arg Val Ala Arg Tyr Thr Thr 1300	1305	1310
Ile Ser Asn Asp Asn Leu Ser Phe Val Ile Ser Asp Lys Lys Val Asp 1315	1320	1325
Thr Asn Phe Ile Tyr Gln Gln Gly Met Leu Leu Gly Leu Gly Val Leu 1330	1335	1340
Glu Thr Leu Phe Arg Leu Glu Lys Asp Thr Gly Ser Ser Asn Thr Val 1345	1350	1355 1360
Leu His Leu His Val Glu Thr Asp Cys Cys Val Ile Pro Met Ile Asp 1365	1370	1375
His Pro Arg Ile Pro Ser Ser Arg Lys Leu Glu Leu Arg Ala Glu Leu 1380	1385	1390
Cys Thr Asn Pro Leu Ile Tyr Asp Asn Ala Pro Leu Ile Asp Arg Asp 1395	1400	1405
Thr Thr Arg Leu Tyr Thr Gln Ser His Arg Arg His Leu Val Glu Phe 1410	1415	1420
Val Thr Trp Ser Thr Pro Gln Leu Tyr His Ile Leu Ala Lys Ser Thr 1425	1430	1435 1440
Ala Leu Ser Met Ile Asp Leu Val Thr Lys Phe Glu Lys Asp His Met 1445	1450	1455
Asn Glu Ile Ser Ala Leu Ile Gly Asp Asp Asp Ile Asn Ser Phe Ile 1460	1465	1470
Thr Glu Phe Leu Leu Ile Glu Pro Arg Leu Phe Thr Ile Tyr Leu Gly 1475	1480	1485
Gln Cys Ala Ala Ile Asn Trp Ala Phe Asp Val His Tyr His Arg Pro 1490	1495	1500
Ser Gly Lys Tyr Gln Met Gly Glu Leu Leu Ser Ser Phe Leu Ser Arg 1505	1510	1515 1520
Met Ser Lys Gly Val Phe Lys Val Leu Val Asn Ala Leu Ser His Pro 1525	1530	1535
Lys Ile Tyr Lys Lys Phe Trp His Cys Gly Ile Ile Glu Pro Ile His 1540	1545	1550
Gly Pro Ser Leu Asp Ala Gln Asn Leu His Thr Thr Val Cys Asn Met 1555	1560	1565

SUBSTITUTE SHEET (RULE 26)

- 234 -

Val Tyr Thr Cys Tyr Met Thr Tyr Leu Asp Leu Leu Leu Asn Glu Glu  
 1570 1575 1580

Leu Glu Glu Phe Thr Phe Leu Leu Cys Glu Ser Asp Glu Asp Val Val  
 1585 1590 1595 1600

Pro Asp Arg Phe Asp Asn Ile Gln Ala Lys His Leu Cys Val Leu Ala  
 1605 1610 1615

Asp Leu Tyr Cys Gln Pro Gly Ala Cys Pro Pro Ile Arg Gly Leu Arg  
 1620 1625 1630

Pro Val Glu Lys Cys Ala Val Leu Thr Asp His Ile Lys Ala Glu Ala  
 1635 1640 1645

Arg Leu Ser Pro Ala Gly Ser Ser Trp Asn Ile Asn Pro Ile Ile Val  
 1650 1655 1660

Asp His Tyr Ser Cys Ser Leu Thr Tyr Leu Arg Arg Gly Ser Ile Lys  
 1665 1670 1675 1680

Gln Ile Arg Leu Arg Val Asp Pro Gly Phe Ile Phe Asp Ala Leu Ala  
 1685 1690 1695

Glu Val Asn Val Ser Gln Pro Lys Ile Gly Ser Asn Asn Ile Ser Asn  
 1700 1705 1710

Met Ser Ile Lys Ala Phe Arg Pro Pro His Asp Asp Val Ala Lys Leu  
 1715 1720 1725

Leu Lys Asp Ile Asn Thr Ser Lys His Asn Leu Pro Ile Ser Gly Gly  
 1730 1735 1740

Asn Leu Ala Asn Tyr Glu Ile His Ala Phe Arg Arg Ile Gly Leu Asn  
 1745 1750 1755 1760

Ser Ser Ala Cys Tyr Lys Ala Val Glu Ile Ser Thr Leu Ile Arg Arg  
 1765 1770 1775

Cys Leu Glu Pro Gly Glu Asp Gly Leu Phe Leu Gly Glu Gly Ser Gly  
 1780 1785 1790

Ser Met Leu Ile Thr Tyr Lys Glu Ile Leu Lys Leu Asn Lys Cys Phe  
 1795 1800 1805

Tyr Asn Ser Gly Val Ser Ala Asn Ser Arg Ser Gly Gln Arg Glu Leu  
 1810 1815 1820

Ala Pro Tyr Pro Ser Glu Val Gly Leu Val Glu His Arg Met Gly Val  
 1825 1830 1835 1840

SUBSTITUTE SHEET (RULE 26)

- 235 -

Gly Asn Ile Val Lys Val Leu Phe Asn Gly Arg Pro Glu Val Thr Trp  
 1845 1850 1855  
 Val Gly Ser Val Asp Cys Phe Asn Phe Ile Val Ser Asn Ile Pro Thr  
 1860 1865 1870  
 Ser Ser Val Gly Phe Ile His Ser Asp Ile Glu Thr Leu Pro Asn Lys  
 1875 1880 1885  
 Asp Thr Ile Glu Lys Leu Glu Glu Leu Ala Ala Ile Leu Ser Met Ala  
 1890 1895 1900  
 Leu Leu Leu Gly Lys Ile Gly Ser Ile Leu Val Ile Lys Leu Met Pro  
 1905 1910 1915 1920  
 Phe Ser Gly Asp Phe Val Gln Gly Phe Ile Ser Tyr Val Gly Ser Tyr  
 1925 1930 1935  
 Tyr Arg Glu Val Asn Leu Val Tyr Pro Arg Tyr Ser Asn Phe Ile Ser  
 1940 1945 1950  
 Thr Glu Ser Tyr Leu Val Met Thr Asp Leu Lys Ala Asn Arg Leu Met  
 1955 1960 1965  
 Asn Pro Glu Lys Ile Lys Gln Gln Ile Ile Glu Ser Ser Val Arg Thr  
 1970 1975 1980  
 Ser Pro Gly Leu Ile Gly His Ile Leu Ser Ile Lys Gln Leu Ser Cys  
 1985 1990 1995 2000  
 Ile Gln Ala Ile Val Gly Asp Ala Val Ser Arg Gly Asp Ile Asn Pro  
 2005 2010 2015  
 Thr Leu Lys Lys Leu Thr Pro Ile Glu Gln Val Leu Ile Asn Cys Gly  
 2020 2025 2030  
 Leu Ala Ile Asn Gly Pro Lys Leu Cys Lys Glu Leu Ile His His Asp  
 2035 2040 2045  
 Val Ala Ser Gly Gln Asp Gly Leu Leu Asn Ser Ile Leu Ile Leu Tyr  
 2050 2055 2060  
 Arg Glu Leu Ala Arg Phe Lys Asp Asn Arg Arg Ser Gln Gln Gly Met  
 2065 2070 2075 2080  
 Phe His Ala Tyr Pro Val Leu Val Ser Ser Arg Gln Arg Glu Leu Ile  
 2085 2090 2095  
 Ser Arg Ile Thr Arg Lys Phe Trp Gly His Ile Leu Leu Tyr Ser Gly  
 2100 2105 2110  
 Asn Arg Lys Leu Ile Asn Lys Phe Ile Gln Asn Leu Lys Ser Gly Tyr

SUBSTITUTE SHEET (RULE 26)

- 236 -

2115                      2120                      2125  
 Leu Ile Leu Asp Leu His Gln Asn Ile Phe Val Lys Asn Leu Ser Lys  
       2130                      2135                      2140  
 Ser Glu Lys Gln Ile Ile Met Thr Gly Gly Leu Lys Arg Glu Trp Val  
       2145                      2150                      2155                      2160  
 Phe Lys Val Thr Val Lys Glu Thr Lys Glu Trp Tyr Lys Leu Val Gly  
                              2165                      2170                      2175  
 Tyr Ser Ala Leu Ile Lys Asp  
                              2180

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15462 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACCAAACAAG AGAAGAACT TGTCTGGGAA TATAAATTTA ACTTTAAATT AACTTAGGAT	60
TAAAGACATT GACTAGAAGG TCAAGAAAAG GGAAGCTCTAT AATTTCAAAA ATGTTGAGCC	120
TATTTGATAC ATTTAATGCA CGTAGGCAAG AAAACATAAC AAAATCAGCC GGTGGAGCTA	180
TCATTCTCTG ACAGAAAAAT ACTGTCTCTA TATTCGCCCT TGGACCGACA ATAAGTGATG	240
ATAATGAGAA AATGACATTA GCTCTTCTAT TTCTATCTCA TTCACTAGAT AATGAGAAAC	300
AACATGCACA AAGGGCAGGG TTCTTGGTGT CTTTATTGTC AATGGCTTAT GCCAATCCAG	360
AGCTCTACCT AACAAACAAAT GGAAGTAATG CAGATGTCAA GTATGTCATA TACATGATTG	420
AGAAAGATCT AAAACGGCAA AAGTATGGAG GATTTGTGGT TAAGACGAGA GAGATGATAT	480
ATGAAAAGAC AACTGATTGG ATATTTGGAA GTGACCTGGA TTATGATCAG GAACTATGT	540
TGCAGAACGG CAGGAACAAT TCAACAATTG AAGACCTTGT CCACACATT GGGTATCCAT	600
CATGTTTAGG AGCTCTTATA ATACAGATCT GGATAGTTCT GGTCAAAGCT ATCACTAGTA	660
TCTCAGGGTT AAGAAAAGGC TTTTTCACCC GATTGGAAGC TTTGAGACAA GATGGAACAG	720

- 237 -

TGCAGGCAGG GCTGGTATTG AGCGGTGACA CAGTGGATCA GATTGGGTCA ATCATGCGGT	780
CTCAACAGAG CTTGGTAACT CTTATGGTTG AAACATTAAT AACAATGAAT ACCAGCAGAA	840
ATGACCTCAC AACCATAGAA AAGAATATAC AAATTGTTGG CAACTACATA AGAGATGCAG	900
GTCTCGCTTC ATTCTTCAAT ACAATCAGAT ATGGAATTGA GACCAGAATG GCAGCTTTGA	960
CTCTATCCAC TCTCAGACCA GATATCAATA GATTAAAAGC TTTGATGGAA CTGTATTTAT	1020
CAAAGGGACC ACGCGCTCCT TTCATCTGTA TCCTCAGAGA TCCTATACAT GGTGAGTTCG	1080
CACCAGGCAA CTATCCTGCC ATATGGAGCT ATGCAATGGG GGTGGCAGTT GTACAAAATA	1140
GAGCCATGCA ACAGTATGTG ACGGGAAGAT CATATCTAGA CATTGATATG TTCCAGCTAG	1200
GACAAGCAGT AGCACGTGAT GCCGAAGCTC AAATGAGCTC AACACTGGAA GATGAACCTG	1260
GAGTGACACA CGAATCTAAA GAAAGCTTGA AGAGACATAT AAGGAACATA AACAGTTCAG	1320
AGACATCTTT CCACAAACCG ACAGGTGGAT CAGCCATAGA GATGGCAATA GATGAAGAGC	1380
CAGAACAATT CGAACATAGA GCAGATCAAG AACAAAATGG AGAACCTCAA TCATCCATAA	1440
TTCAATATGC CTGGGCAGAA GGAAATAGAA GCGATGATCA GACTGAGCAA GCTACAGAAT	1500
CTGACAATAT CAAGACCGAA CAACAAAACA TCAGAGACAG ACTAAACAAG AGACTCAACG	1560
ACAAGAAGAA ACAAAGCAGT CAACCACCCA CTAATCCCAC AAACAGAACA AACCAGGACG	1620
AAATAGATGA TCTGTTTAAAC GCATTTGGAA GCAACTAATC GAATCAACAT TTTAATCTAA	1680
ATCAATAATA AATAAGAAAA ACTTAGGATT AAAGAATCCT ATCATACCGG AATATAGGGT	1740
GGTAAATTTA GAGTCTGCTT GAAACTCAAT CAATAGAGAG TTGATGGAAA GCGATGCTAA	1800
AAACTATCAA ATCATGGATT CTTGGGAAGA GGAATCAAGA GATAAATCAA CTAATATCTC	1860
CTCGGCCCTC AACATCATTG AATTCATACT CAGCACCGAC CCCCAGAAG ACTTATCGGA	1920
AAACGACACA ATCAACACAA GAACCCAGCA ACTCAGTGCC ACCATCTGTC AACCAGAAAT	1980
CAAACCAACA GAAACAAGTG AGAAAGATAG TGGATCAACT GACAAAAATA GACAGTCCGG	2040
GTCATCACAC GAATGTACAA CAGAAGCAAA AGATAGAAAT ATTGATCAGG AAAGTGTACA	2100
GAGAGGACCT GGGAGAAGAA GCAGCTCAGA TAGTAGAGCT GAGACTGTGG TCTCTGGAGG	2160
AATCCCCAGA AGCATCACAG ATTCTAAAAA TGGAACCCAA AACACGGAGG ATATTGATCT	2220
CAATGAAATT AGAAAGATGG ATAAGGACTC TATTGAGGGG AAAATGCGAC AATCTGCAAA	2280

SUBSTITUTE SHEET (RULE 26)

- 238 -

TGTTCCAAGC GAGATATCAG GAAGTGATGA CATATTTACA ACAGAACAAA GTAGAAACAG	2340
TGATCATGGA AGAAGCCTGG AATCTATCAG TACACCTGAT ACAAGATCAA TAAGTGTGT	2400
TACTGCTGCA ACACCAGATG ATGAAGAAGA AATACTAATG AAAAATAGTA GGACAAAGAA	2460
AAGTTCTTCA ACACATCAAG AAGATGACAA AAGAATTAAA AAAGGGGGAA AAGGGAAAGA	2520
CTGGTTTAAG AAATCAAAAG ATACCGACAA CCAGATACCA ACATCAGACT ACAGATCCAC	2580
ATCAAAAGGG CAGAAGAAAA TCTCAAAGAC AACCAACCACC AACACCGACA CAAAGGGGCA	2640
AACAGAAATA CAGACAGAAT CATCAGAAAC ACAATCCTCA TCATGGAATC TCATCATCGA	2700
CAACAACACC GACCGGAACG AACAGACAAG CACAACCTCT CCAACAACAA CTTCCAGATC	2760
AACTTATACA AAAGAATCGA TCCGAACAAA CTCTGAATCC AAACCCAAGA CACAAAAGAC	2820
AAATGGAAAG GAAAGGAAGG ATACAGAAGA GAGCAATCGA TTTACAGAGA GGGCAATTAC	2880
TCTATTGCAG AATCTTGGTG TAATTCAATC CACATCAAAA CTAGATTTAT ATCAAGACAA	2940
ACGAGTTGTA TGTGTAGCAA ATGTACTAAA CAATGTAGAT ACTGCATCAA AGATAGATTT	3000
CCTGGCAGGA TTAGTCATAG GGGTTTCAAT GGACAACGAC ACAAATTA CACAGATACA	3060
AAATGAAATG CTAAACCTCA AAGCAGATCT AAAGAAAATG GACGAATCAC ATAGAAGATT	3120
GATAGAAAAT CAAAGAGAAC AACTGTCATT GATCACGTCA CTAATTTCAA ATCTCAAAAT	3180
TATGACTGAG AGAGGAGGAA AGAAAGACCA AAATGAATCC AATGAGAGAG TATCCATGAT	3240
CAAAACAAAA TTGAAAGAAG AAAAGATCAA GAAGACCAGG TTTGACCCAC TTATGGAGGC	3300
ACAAGGCATT GACAAGAATA TACCCGATCT ATATCGACAT GCAGGAGATA CACTAGAGAA	3360
CGATGTACAA GTTAAATCAG AGATATTAAG TTCATACAAT GAGTCAAATG CAACAAGACT	3420
AATACCCAAA AAAGTGAGCA GTACAATGAG ATCACTAGTT GCAGTCATCA ACAACAGCAA	3480
TCTCTACAA AGCACAAAAC AATCATACAT AAACGAACTC AAACGTTGCA AAAATGATGA	3540
AGAAGTATCT GAATTAATGG ACATGTTCAA TGAAGATGTC AACAATTGCC AATGATCCAA	3600
CAAAGAAACG ACACCGAACA AACAGACAAG AAACAACAGT AGATCAAAAC CTGTCAACAC	3660
ACACAAAATC AAGCAGAATG AAACAACAGA TATCAATCAA TATACAAATA AGAAAACTT	3720
AGGATTAAAG AATAAATTAA TCCTTGTTCA AAATGAGTAT AACTAACTCT GCAATATACA	3780
CATTCCGAGA ATCATCATT TCTGAAAATG GTCATATAGA ACCATTACCA CTCAAAGTCA	3840

SUBSTITUTE SHEET (RULE 26)

- 239 -

ATGAACAGAG GAAAGCAGTA CCCCACATTA GAGTTGCCAA GATCGGAAAT CCACCAAAAC	3900
ACGGATCCCG GTATTTAGAT GTCTTCTTAC TCGGCTTCTT CGAGATGGAA CGAATCAAAG	3960
ACAAATACGG GAGTGTGAAT GATCTCGACA GTGACCCGAG TTACAAAGTT TGTGGCTCTG	4020
GATCATTACC AATCGGATTG GCTAAGTACA CTGGGAATGA CCAGGAATTG TTACAAGCCG	4080
CAACCAAACT GGATATAGAA GTGAGAAGAA CAGTCAAAGC GAAAGAGATG GTTGTTTACA	4140
CGGTACAAAA TATAAAACCA GAACTGTACC CATGGTCCAA TAGACTAAGA AAAGGAATGC	4200
TGTTTCGATGC CAACAAAGTT GCTCTTGCTC CTCAATGTCT TCCACTAGAT AGGAGCATAA	4260
AATTTAGAGT AATCTTCGTG AATTGTACGG CAATTGGATC AATAACCTTG TTCAAATTC	4320
CTAAGTCAAT GGCATCACTA TCTCTACCCA ACACAATATC AATCAATCTG CAGGTACACA	4380
TAAAAACAGG GGTTCAGACT GATTCTAAAG GGATAGTTCA AATTTTGGAT GAGAAAGGCG	4440
AAAAATCACT GAATTTTCATG GTCCATCTCG GATTGATCAA AAGAAAAGTA GGCAGAATGT	4500
ACTCTGTTGA ATACTGTAAA CAGAAAATCG AGAAAATGAG ATTGATATTT TCTTTAGGAC	4560
TAGTTGGAGG AATCAGTCTT CATGTCAATG CAACTGGGTC CATATCAAAA AACTAGCAA	4620
GTCAGCTGGT ATTCAAAAGA GAGATTTGTT ATCCTTTAAT GGATCTAAAT CCGCATCTCA	4680
ATCTAGTTAT CTGGGCTTCA TCAGTAGAGA TTACAAGAGT GGATGCAATT TTCCAACCTT	4740
CTTTACCTGG CGAGTTCAGA TACTATCCTA ATATTATTGC AAAAGGAGTT GGGAAAATCA	4800
AACAATGGAA CTAGTAATCT CTATTTTAGT CCGGACGTAT CTATTAAGCC GAAGCAAATA	4860
AAGGATAATC AAAAAGTTAG GACAAAAGAG GTCAATACCA ACAACTATTA GCAGTCACAC	4920
TCGCAAGAAT AAGAGAGAAG GGACCAAAAA AGTCAAATAG GAGAAATCAA AACAAAAGGT	4980
ACAGAACACC AGAACAACAA AATCAAAACA TCCAATCAC TCAAAACAAA AATTCCAAAA	5040
GAGACCGGCA ACACAACAAG CACTGAACAC AATGCCAACT TCAATACTGC TAATTATTAC	5100
AACCATGATC ATGGCATCTT TCTGCCAAAT AGATATCACA AACTACAGC ACGTAGGTGT	5160
ATTGGTCAAC AGTCCCAAAG GGATGAAGAT ATCACAAAAC TTTGAAACAA GATATCTAAT	5220
TTTGAGCCTC ATACCAAAAA TAGAAGACTC TAACTCTTGT GGTGACCAAC AGATCAAGCA	5280
ATACAAGAAG TTATTGGATA GACTGATCAT CCCTTTATAT GATGGATTAA GATTACAGAA	5340
AGATGTGATA GTAACCAATC AAGAATCCAA TGAAAACACT GATCCCAGAA CAAAACGATT	5400

SUBSTITUTE SHEET (RULE 26)

- 240 -

CTTTGGAGGG	GTAATTGGAA	CCATTGCTCT	GGGAGTAGCA	ACCTCAGCAC	AAATTACAGC	5460
GGCAGTTGCT	CTGGTTGAAG	CCAAGCAGGC	AAGATCAGAC	ATCGAAAAAC	TCAAAGAAGC	5520
AATTAGGGAC	ACAAACAAAG	CAGTGCAGTC	AGTTCAGAGC	TCCATAGGAA	ATTTAATAGT	5580
AGCAATTAAA	TCAGTCCAGG	ATTATGTTAA	CAAAGAAATC	GTGCCATCGA	TGCGAGGGCT	5640
AGGTTGTGAA	GCAGCAGGAC	TTCAATTAGG	AATTGCATTA	ACACAGCATT	ACTCAGAATT	5700
AACAAACATA	TTTGGTGATA	ACATAGGATC	GTTACAAGAA	AAAGGAATAA	AATTACAAGG	5760
TATAGCATCA	TTATACCGCA	CAAATATCAC	AGAAATATTC	ACAACATCAA	CAGTTGATAA	5820
ATATGATATC	TATGATCTGT	TATTTACAGA	ATCAATAAAG	GTGAGAGTTA	TAGATGTTGA	5880
CTTGAATGAT	TACTCAATCA	CCCTCCAAGT	CAGACTCCCT	TTATTAACTA	GGCTGCTGAA	5940
CACTCAGATC	TACAAAGTAG	ATTCCATATC	ATATAACATC	CAAAACAGAG	AATGGTATAT	6000
CCCTCTTCCC	AGCCATATCA	TGACGAAAGG	GGCATTCTTA	GGTGGAGCAG	ACGTCAAAGA	6060
ATGTATAGAA	GCATTCAGCA	GCTATATATG	CCCTTCTGAT	CCAGGATTTG	TATTAAACCA	6120
TGAAATAGAG	AGCTGCTTAT	CAGGAAACAT	ATCCCAATGT	CCAAGAACAA	CGGTCACATC	6180
AGACATTGTT	CCAAGATATG	CATTTGTCAA	TGGAGGAGTG	GTTGCAAACT	GTATAACAAC	6240
CACCTGTACA	TGCAACGGAA	TTGGTAATAG	AATCAATCAA	CCACCTGATC	AAGGAGTAAA	6300
AATTATAACA	CATAAAGAAT	GTAGTACAAT	AGGTATCAAC	GGAATGCTGT	TCAATACAAA	6360
TAAAGAAGGA	ACTCTTGCA	TCTATACACC	AAATGATATA	ACACTAAACA	ATTCTGTTGC	6420
ACTTGATCCA	ATTGACATAT	CAATCGAGCT	CAACAAGGCC	AAATCAGATC	TAGAAGAATC	6480
AAAAGAATGG	ATAAGAAGGT	CAAATCAAAA	ACTAGATTCT	ATTGGAAATT	GGCATCAATC	6540
TAGCACTACA	ATCATAATTA	TTTTGATAAT	GATCATTATA	TTGTTTATAA	TTAATATAAC	6600
GATAATTACA	ATTGCAATTA	AGTATTACAG	AATTCAAAAG	AGAAATCGAG	TGGATCAAAA	6660
TGACAAGCCA	TATGTACTAA	CAAACAAATA	ACATATCTAC	AGATCATTAG	ATATTAAAAT	6720
TATAAAAAAC	TTAGGAGTAA	AGTTACGCAA	TCCAACCTCTA	CTCATATAAT	TGAGGAAGGA	6780
CCCAATAGAC	AAATCCAAAT	TCGAGATGGA	ATACTGGAAG	CATACCAATC	ACGGAAAGGA	6840
TGCTGGTAAT	GAGCTGGAGA	CGTCTATGGC	TACTCATGGC	AACAAGCTCA	CTAATAAGAT	6900
AATATACATA	TTATGGACAA	TAATCCTGGT	GTTATTATCA	ATAGTCTTCA	TCATAGTGCT	6960

SUBSTITUTE SHEET (RULE 26)



- 241 -

AATTAATTCC ATCAAAAGTG AAAAGGCCCA CGAATCATTG CTGCAAGACA TAAATAATGA 7020  
GTTTATGGAA ATTACAGAAA AGATCCAAAT GGCATCGGAT AATACCAATG ATCTAATACA 7080  
GTCAGGAGTG AATACAAGGC TTCTTACAAT TCAGAGTCAT GTCCAGAATT ACATACCAAT 7140  
ATCATTGACA CAACAGATGT CAGATCTTAG GAAATTCATT AGTGAAATTA CAATTAGAAA 7200  
TGATAATCAA GAAGTGCTGC CACAAAGAAT AACACATGAT GTAGGTATAA AACCTTTAAA 7260  
TCCAGATGAT TTTTGGAGAT GCACGTCTGG TCTTCCATCT TTAATGAAAA CTCCAAAAAT 7320  
AAGGTTAATG CCAGGGCCGG GATTATTAGC TATGCCAACG ACTGTTGATG GCTGTGTTAG 7380  
AACTCCGTCT TTAGTTATAA ATGATCTGAT TTATGCTTAT ACCTCAAATC TAATTACTCG 7440  
AGGTTGTCAG GATATAGGAA AATCATATCA AGTCTTACAG ATAGGGATAA TAACTGTAAA 7500  
CTCAGACTTG GTACCTGACT TAAATCCTAG GATCTCTCAT ACCTTTAACA TAAATGACAA 7560  
TAGGAAGTCA TGTCTCTAG CACTCCTAAA TACAGATGTA TATCAACTGT GTTCAACTCC 7620  
CAAAGTTGAT GAAAGATCAG ATTATGCATC ATCAGGCATA GAAGATATTG TACTTGATAT 7680  
TGTCAATTAT GATGGTTCAA TCTCAACAAC AAGATTTAAG AATAATAACA TAAGCTTTGA 7740  
TCAACCATAT GCTGCACTAT ACCCATCTGT TGGACCAGGG ATATACTACA AAGGCAAAAT 7800  
AATATTTCTC GGGTATGGAG GTCTTGAACA TCCAATAAAT GAGAATGTAA TCTGCAACAC 7860  
AACTGGGTGC CCCGGGAAAA CACAGAGAGA CTGTAATCAA GCGTCTCATA GTCCATGGTT 7920  
TTCAGATAGG AGGATGGTCA ACTCCATCAT TGTGTTGAC AAAGGCTTAA ACTCAATTCC 7980  
AAAATTGAAA GTATGGACGA TATCTATGCG ACAAATTAC TGGGGGTCAG AAGGAAGGTT 8040  
ACTTCTACTA GGTAACAAGA TCTATATATA TACAAGATCT ACAAGTTGGC ATAGCAAGTT 8100  
ACAATTAGGA ATAATTGATA TTAAGTATTA CAGTGATATA AGGATAAAAT GGACATGGCA 8160  
TAATGTGCTA TCAAGACCAG GAAACAATGA ATGTCCATGG GGACATTCAT GTCCAGATGG 8220  
ATGTATAACA GGAGTATATA CTGATGCATA TCCACTCAAT CCCACAGGGA GCATTGTGTC 8280  
ATCTGTCATA TTAGACTCAC AAAAATCGAG AGTGAACCCA GTCATAACTT ACTCAACAGC 8340  
AACCAGAAAG GTAAACGAGC TGGCCATCCT AAACAGAACA CTCTCAGCTG GATATACAAC 8400  
AACAGCTGCT ATTACACACT ATAACAAAGG ATATTGTTTT CATATAGTAG AAATAAATCA 8460  
TAAAAGCTTA AACACATTTT AACCCATGTT GTTCAAAACA GAGATTCCAA AAAGCTGCAG 8520

SUBSTITUTE SHEET (RULE 26)

- 242 -

TTAATCATAA TTAACCATAA TATGCATCAA TCTATCTATA ATACAAGTAT ATGATAAGTA	8580
ATCAGCAATC AGACAATAGA CAAAAGGGAA ATATAAAAAA CTTAGGAGCA AAGCGTGCTC	8640
GGGAAATGGA CACTGAATCT AACAAATGGCA CTGTATCTGA CATACTCTAT CCTGAGTGTC	8700
ACCTTAACTC TCCTATCGTT AAAGGTAAAA TAGCACAATT ACACACTATT ATGAGTCTAC	8760
CTCAGCCTTA TGATATGGAT GACGACTCAA TACTAGTTAT CACTAGACAG AAAATAAAAC	8820
TTAATAAATT GGATAAAAGA CAACGATCTA TTAGAAGATT AAAATTAATA TTAAGTAA	8880
AAGTGAATGA CTTAGGAAAA TACACATTTA TCAGATATCC AGAAATGTCA AAAGAAATGT	8940
TCAAATTATA TATACCTGGT ATTAACAGTA AAGTGACTGA ATTATTACTT AAAGCAGATA	9000
GAACATATAG TCAAATGACT GATGGATTAA GAGATCTATG GATTAATGTG CTATCAAAAT	9060
TAGCCTCAA AAATGATGGA AGCAATTATG ATCTTAATGA AGAAATTAAT AATATATCGA	9120
AAGTTCACAC AACCTATAAA TCAGATAAAT GGTATAATCC ATTCAAAACA TGGTTTACTA	9180
TCAAGTATGA TATGAGAAGA TTACAAAAAG CTCGAAATGA GATCACTTTT AATGTTGGGA	9240
AGGATTATAA CTTGTTAGAA GACCAGAAGA ATTTCTTATT GATACATCCA GAATTGGTTT	9300
TGATATTAGA TAAACAAAAC TATAATGGTT ATCTAATTAC TCCTGAATTA GTATTGATGT	9360
ATTGTGACGT AGTCGAAGGC CGATGGAATA TAAGTGCATG TGCTAAGTTA GATCCAAAAT	9420
TACAATCTAT GTATCAGAAA GGTAATAACC TGTGGGAAGT GATAGATAAA TTGTTTCCAA	9480
TTATGGGAGA AAAGACATTT GATGTGATAT CGTTATTAGA ACCACTTGCA TTATCCTTAA	9540
TTCAAACCTCA TGATCCTGTT AAACAACATA GAGGAGCTTT TTAAATCAT GTGTTATCCG	9600
AGATGGAATT AATATTTGAA TCTAGAGAAT CGATTAAGGA ATTTCTGAGT GTAGATTACA	9660
TTGATAAAAT TTTAGATATA TTTAATAAGT CTACAATAGA TGAAATAGCA GAGATTTTCT	9720
CTTTTTTTAG AACATTTGGG CATCCTCCAT TAGAAGCTAG TATTGCAGCA GAAAAGGTTA	9780
GAAAATATAT GTATATTGGA AAACAATTAA AATTGACAC TATTAATAAA TGTCATGCTA	9840
TCTTCTGTAC AATAATAATT AACGGATATA GAGAGAGGCA TGGTGGACAG TGGCCTCCTG	9900
TGACATTACC TGATCATGCA CACGAATTCA TCATAAATGC TTACGGTTCA AACTCTGCCA	9960
TATCATATGA AAATGCTGTT GATTATTACC AGAGCTTTAT AGGAATAAAA TTCAATAAAT	10020
TCATAGAGCC TCAGTTAGAT GAGGATTTGA CAATTTATAT GAAAGATAAA GCATTATCTC	10080

SUBSTITUTE SHEET (RULE 26)

- 243 -

CAAAAAAATC	AAATTGGGAC	ACAGTTTATC	CTGCATCTAA	TTTACTGTAC	CGTACTAACG	10140
CATCCAACGA	ATCACGAAGA	TTAGTTGAAG	TATTTATAGC	AGATAGTAAA	TTTGATCCTC	10200
ATCAGATATT	GGATTATGTA	GAATCTGGGG	ACTGGTTAGA	TGATCCAGAA	TTTAATATTT	10260
CTTATAGTCT	TAAAGAAAAA	GAGATCAAAC	AGGAAGGTAG	ACTCTTTGCA	AAAATGACAT	10320
ACAAAATGAG	AGCTACACAA	GTTTTATCAG	AGACACTACT	TGCAAATAAC	ATAGGAAAAT	10380
TCTTTCAAGA	AAATGGGATG	GTGAAGGGAG	AGATTGAATT	ACTTAAGAGA	TTAACAACCA	10440
TATCAATATC	AGGAGTTCCA	CGGTATAATG	AAGTGTACAA	TAATTCTAAA	AGCCATACAG	10500
ATGACCTTAA	AACCTACAAT	AAAATAAGTA	ATCTTAATTT	GTCTTCTAAT	CAGAAATCAA	10560
AGAAATTTGA	ATTCAAGTCA	ACGGATATCT	ACAATGATGG	ATACGAGACT	GTGAGCTGTT	10620
TCCTAACAAC	AGATCTCAAA	AAATACTGTC	TTAATTGGAG	ATATGAATCA	ACAGCTCTAT	10680
TTGGAGAAAC	TTGCAACCAA	ATATTTGGAT	TAAATAAATT	GTTTAATTGG	TTACACCCTC	10740
GTCTTGAAGG	AAGTACAATC	TATGTAGGTG	ATCCTTACTG	TCCTCCATCA	GATAAAGAAC	10800
ATATATCATT	AGAGGATCAC	CCTGATTCTG	GTTTTTACGT	TCATAACCCA	AGAGGGGGTA	10860
TAGAAGGATT	TTGTCAAAAA	TTATGGACAC	TCATATCTAT	AAGTGCAATA	CATCTAGCAG	10920
CTGTTAGAAT	AGGCGTGAGG	GTGACTGCAA	TGGTTCAAGG	AGACAATCAA	GCTATAGCTG	10980
TAACCACAAG	AGTACCCAAC	AATTATGACT	ACAGAGTTAA	GAAGGAGATA	GTTTATAAAG	11040
ATGTAGTGAG	ATTTTTTGAT	TCATTAAGAG	AAGTGATGGA	TGATCTAGGT	CATGAACCTA	11100
AATTAATGA	AACGATTATA	AGTAGCAAGA	TGTTCAATATA	TAGCAAAAGA	ATCTATTATG	11160
ATGGGAGAAT	TCTTCCTCAA	GCTCTAAAAG	CATTATCTAG	ATGTGTCTTC	TGGTCAGAGA	11220
CAGTAATAGA	CGAAACAAGA	TCAGCATCTT	CAAATTTGGC	AACATCATTT	GCAAAAGCAA	11280
TTGAGAATGG	TTATTCACCT	GTTCTAGGAT	ATGCATGCTC	AATTTTAAAG	AACATTCAAC	11340
AACTATATAT	TGCCCTTGGG	ATGAATATCA	ATCCAACAT	AACACAGAAT	ATCAGAGATC	11400
AGTATTTTAG	GAATCCAAAT	TGGATGCAAT	ATGCCTCTTT	AATACCTGCT	AGTGTTGGGG	11460
GATTCAATTA	CATGGCCATG	TCAAGATGTT	TTGTAAGGAA	TATTGGTGAT	CCATCAGTTG	11520
CCGCATTGGC	TGATATTAAA	AGATTTATTA	AGGCGAATCT	ATTAGACCGA	AGTGTTCTTT	11580
ATAGGATTAT	GAATCAAGAA	CCAGGTGAGT	CATCTTTTTT	GGACTGGGCT	TCAGATCCAT	11640

SUBSTITUTE SHEET (RULE 26)

- 244 -

ATTCATGCAA	TTTACCACAA	TCTCAAAATA	TAACCACCAT	GATAAAAAAT	ATAACAGCAA	11700
GGAAATGTATT	ACAAGATTCA	CCAAATCCAT	TATTATCTGG	ATTATTCACA	AATACAATGA	11760
TAGAAGAAGA	TGAAGAATTA	GCTGAGTTCC	TGATGGACAG	GAAGGTAATT	CTCCCTAGAG	11820
TTGCACATGA	TATTCTAGAT	AATTCTCTCA	CAGGAATTAG	AAATGCCATA	GCTGGAATGT	11880
TAGATACGAC	AAAATCACTA	ATTCGGGTTG	GCATAAATAG	AGGAGGACTG	ACATATAGTT	11940
TGTTGAGGAA	AATCAGTAAT	TACGATCTAG	TACAATATGA	AACACTAAGT	AGGACTTTGC	12000
GACTAATTGT	AAGTGATAAA	ATCAAGTATG	AAGATATGTG	TTCGGTAGAC	CTTGCCATAG	12060
CATTGCGACA	AAAGATGTGG	ATTCATTTAT	CAGGAGGAAG	GATGATAAGT	GGACTTGAAA	12120
CGCCTGACCC	ATTAGAATTA	CTATCTGGGG	TAGTAATAAC	AGGATCAGAA	CATTGTAAAA	12180
TATGTTATTC	TTCAGATGGC	ACAAACCCAT	ATACTTGGAT	GTATTTACCC	GGTAATATCA	12240
AAATAGGATC	AGCAGAAACA	GGTATATCGT	CATTAAGAGT	TCCTTATTTT	GGATCAGTCA	12300
CTGATGAAAG	ATCTGAAGCA	CAATTAGGAT	ATATCAAGAA	TCTTAGTAAA	CCTGCAAAAG	12360
CCGCAATAAG	AATAGCAATG	ATATATACAT	GGGCATTTGG	TAATGATGAG	ATATCTTGGA	12420
TGGAAGCCTC	ACAGATAGCA	CAAACACGTG	CAAATTTTAC	ACTAGATAGT	CTCAAAATTT	12480
TAACACCGGT	AGCTACATCA	ACAAATTTAT	CACACAGATT	AAAGGATACT	GCAACTCAGA	12540
TGAAATTCTC	CAGTACATCA	TTGATCAGAG	TCAGCAGATT	CATAACAATG	TCCAATGATA	12600
ACATGTCTAT	CAAAGAAGCT	AATGAAACCA	AAGATACTAA	TCTTATTTAT	CAACAAATAA	12660
TGTTAACAGG	ATTAAGTGTT	TTCGAATATT	TATTTAGATT	AAAAGAAACC	ACAGGACACA	12720
ACCCTATAGT	TATGCATCTG	CACATAGAAG	ATGAGTGTTG	TATTAAAGAA	AGTTTTAATG	12780
ATGAACATAT	TAATCCAGAG	TCTACATTAG	AATTAATTCTG	ATATCCTGAA	AGTAATGAAT	12840
TTATTTATGA	TAAAGACCCA	CTCAAAGATG	TGGACTTATC	AAAACCTATG	GTTATTAAAG	12900
ACCATTCTTA	CACAATTGAT	ATGAATTATT	GGGATGATAC	TGACATCATA	CATGCAATTT	12960
CAATATGTAC	TGCAATTACA	ATAGCAGATA	CTATGTCACA	ATTAGATCGA	GATAATTTAA	13020
AAGAGATAAT	AGTTATTGCA	AATGATGATG	ATATTAATAG	CTTAATCACT	GAATTTTGA	13080
CTCTTGACAT	ACTTGATTTT	CTCAAGACAT	TTGGTGGATT	ATTAGTAAAT	CAATTGTCAT	13140
ACACTCTTTA	TAGTCTAAAA	ATAGAAGGTA	GGGATCTCAT	TTGGGATTAT	ATAATGAGAA	13200

SUBSTITUTE SHEET (RULE 26)

- 245 -

CACTGAGAGA TACTTCCCAT TCAATATTAA AAGTATTATC TAATGCATTA TCTCATCCTA 13260  
AAGTATTCAA GAGGTTCTGG GATTGTGGAG TTTTAAACCC TATTTATGGT CCTAATACTG 13320  
CTAGTCAAGA CCAGATAAAA CTGCCCCTAT CTATATGTGA ATATTCATA GATCTATTTA 13380  
TGAGAGAATG GTTGAATGGT GTATCACTTG AAATATACAT TTGTGACAGC GATATGGAAG 13440  
TTGCAAATGA TAGGAAACAA GCCTTTATTT CTAGACACCT TTCATTTGTT TGTGTTTAG 13500  
CAGAAATTGC ATCTTTCGGA CCTAACCTGT TAACTTAAAC ATACTTGGAG AGACTTGATC 13560  
TATTGAAACA ATATCTTGAA TTAAATATTA AAGAAGACCC TACTCTTAAA TATGTACAAA 13620  
TATCTGGATT ATTAATTAAA TCGTTCCCAT CAACTGTAAC ATACGTAAGA AAGACTGCAA 13680  
TCAAATATCT AAGGATTGCG GGTATTAGTC CACCTGAGGT AATTGATGAT TGGGATCCGG 13740  
TAGAAGATGA AAATATGCTG GATAACATTG TCAAACTAT AAATGATAAC TGTAATAAAG 13800  
ATAATAAAGG GAATAAAATT AACAAATTCT GGGGACTAGC ACTTAAGAAC TATCAAGTCC 13860  
TTAAATCAG ATCTATAACA AGTGATTCTG ATGATAATGA TAGACTAGAT GCTAATACAA 13920  
GTGGTTTGAC ACTTCCTCAA GGAGGGAATT ATCTATCGCA TCAATTGAGA TTATTCGGAA 13980  
TCAACAGCAC TAGTTGTCTG AAAGCTCTTG AGTTATCACA AATTTTAATG AAGGAAGTCA 14040  
ATAAAGACAA GGACAGGCTC TTCCTGGGAG AAGGAGCAGG AGCTATGCTA GCATGTTATG 14100  
ATGCCACATT AGGACCTGCA GTTAATTATT ATAATTCAGG TTTGAATATA ACAGATGTAA 14160  
TTGGTCAACG AGAATTGAAA ATATTTCCCT CAGAGGTATC ATTAGTAGGT AAAAAATTAG 14220  
GAAATGTGAC ACAGATTCTT AACAGGGTAA AAGTACTGTT CAATGGGAAT CCTAATTCAA 14280  
CATGGATAGG AAATATGGAA TGTGAGAGCT TAATATGGAG TGAATTAAAT GATAAGTCCA 14340  
TTGGATTAGT ACATTGTGAT ATGGAAGGAG CTATCGGTAA ATCAGAAGAA ACTGTTCTAC 14400  
ATGAACATTA TAGTGTTATA AGAATTACAT ACTTGATTGG GGATGATGAT GTTGTTTTAG 14460  
TTTCCAAAAT TATACCTACA ATCACTCCGA ATTGGTCTAG AATACTTTAT CTATATAAAT 14520  
TATATTGGAA AGATGTAAGT ATAATATCAC TCAAACTTC TAATCCTGCA TCAACAGAAT 14580  
TATATCTAAT TTCGAAAGAT GCATATTGTA CTATAATGGA ACCTAGTGAA ATTGTTTTAT 14640  
CAAACTTAA AAGATTGTCA CTCTTGAAG AAAATAATCT ATTAAATGG ATCATTTTAT 14700  
CAAAGAAGAG GAATAATGAA TGGTTACATC ATGAAATCAA AGAAGGAGAA AGAGATTATG 14760

SUBSTITUTE SHEET (RULE 26)

GAATCATGAG	ACCATATCAT	ATGGCACTAC	AAATCTTTGG	ATTTCAAATC	AATTTAAATC	14820
ATCTGGCGAA	AGAATTTTTA	TCAACCCCGAG	ATCTGACTAA	TATCAACAAT	ATAATCCAAA	14880
GTTTTTCAGCG	AACAATAAAG	GATGTTTTTAT	TTGAATGGAT	TAATATAACT	CATGATGATA	14940
AGAGACATAA	ATTAGGCGGA	AGATATAACA	TATTTCCCACT	GAAAAATAAG	GGAAAGTTAA	15000
GACTGCTATC	GAGAAGACTA	GTATTAAGTT	GGATTTCATT	ATCATTATCG	ACTCGATTAC	15060
TTACAGGTCG	CTTTCCTGAT	GAAAAATTTG	AACATAGAGC	ACAGACTGGA	TATGTATCAT	15120
TAGCTGATAC	TGATTTAGAA	TCATTAAAGT	TATTGTCGAA	AAACATCATT	AAGAATTACA	15180
GAGAGTGTAT	AGGATCAATA	TCATATTGGT	TTCTAACCAA	AGAAGTTAAA	ATACTTATGA	15240
AATTGATTGG	TGGTGCTAAA	TTATTAGGAA	TTCCAGACA	ATATAAAGAA	CCCGAAGACC	15300
AGTTATTAGA	AAACTACAAT	CAACATGATG	AATTTGATAT	CGATTAAAAC	ATAAATACAA	15360
TGAAGATATA	TCCTAACCTT	TATCTTTAAG	CCTAGGAATA	GACAAAAAGT	AAGAAAAACA	15420
TGTAATATAT	ATATACCAAA	CAGAGTTCTT	CTCTTGTTTG	GT		15462

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2233 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Thr Glu Ser Asn Asn Gly Thr Val Ser Asp Ile Leu Tyr Pro  
1 5 10 15

Glu Cys His Leu Asn Ser Pro Ile Val Lys Gly Lys Ile Ala Gln Leu  
20 25 30

His Thr Ile Met Ser Leu Pro Gln Pro Tyr Asp Met Asp Asp Asp Ser  
35 40 45

Ile Leu Val Ile Thr Arg Gln Lys Ile Lys Leu Asn Lys Leu Asp Lys  
50 55 60

- 247 -

Arg Gln Arg Ser Ile Arg Arg Leu Lys Leu Ile Leu Thr Glu Lys Val  
 65 70 75 80  
 Asn Asp Leu Gly Lys Tyr Thr Phe Ile Arg Tyr Pro Glu Met Ser Lys  
 85 90 95  
 Glu Met Phe Lys Leu Tyr Ile Pro Gly Ile Asn Ser Lys Val Thr Glu  
 100 105 110  
 Leu Leu Leu Lys Ala Asp Arg Thr Tyr Ser Gln Met Thr Asp Gly Leu  
 115 120 125  
 Arg Asp Leu Trp Ile Asn Val Leu Ser Lys Leu Ala Ser Lys Asn Asp  
 130 135 140  
 Gly Ser Asn Tyr Asp Leu Asn Glu Glu Ile Asn Asn Ile Ser Lys Val  
 145 150 155 160  
 His Thr Thr Tyr Lys Ser Asp Lys Trp Tyr Asn Pro Phe Lys Thr Trp  
 165 170 175  
 Phe Thr Ile Lys Tyr Asp Met Arg Arg Leu Gln Lys Ala Arg Asn Glu  
 180 185 190  
 Ile Thr Phe Asn Val Gly Lys Asp Tyr Asn Leu Leu Glu Asp Gln Lys  
 195 200 205  
 Asn Phe Leu Leu Ile His Pro Glu Leu Val Leu Ile Leu Asp Lys Gln  
 210 215 220  
 Asn Tyr Asn Gly Tyr Leu Ile Thr Pro Glu Leu Val Leu Met Tyr Cys  
 225 230 235 240  
 Asp Val Val Glu Gly Arg Trp Asn Ile Ser Ala Cys Ala Lys Leu Asp  
 245 250 255  
 Pro Lys Leu Gln Ser Met Tyr Gln Lys Gly Asn Asn Leu Trp Glu Val  
 260 265 270  
 Ile Asp Lys Leu Phe Pro Ile Met Gly Glu Lys Thr Phe Asp Val Ile  
 275 280 285  
 Ser Leu Leu Glu Pro Leu Ala Leu Ser Leu Ile Gln Thr His Asp Pro  
 290 295 300  
 Val Lys Gln Leu Arg Gly Ala Phe Leu Asn His Val Leu Ser Glu Met  
 305 310 315 320  
 Glu Leu Ile Phe Glu Ser Arg Glu Ser Ile Lys Glu Phe Leu Ser Val  
 325 330 335  
 Asp Tyr Ile Asp Lys Ile Leu Asp Ile Phe Asn Lys Ser Thr Ile Asp

SUBSTITUTE SHEET (RULE 26)

- 248 -

340	345	350
Glu Ile Ala Glu Ile Phe Ser Phe Phe Arg Thr Phe Gly His Pro Pro		
355	360	365
Leu Glu Ala Ser Ile Ala Ala Glu Lys Val Arg Lys Tyr Met Tyr Ile		
370	375	380
Gly Lys Gln Leu Lys Phe Asp Thr Ile Asn Lys Cys His Ala Ile Phe		
385	390	395
Cys Thr Ile Ile Ile Asn Gly Tyr Arg Glu Arg His Gly Gly Gln Trp		
405	410	415
Pro Pro Val Thr Leu Pro Asp His Ala His Glu Phe Ile Ile Asn Ala		
420	425	430
Tyr Gly Ser Asn Ser Ala Ile Ser Tyr Glu Asn Ala Val Asp Tyr Tyr		
435	440	445
Gln Ser Phe Ile Gly Ile Lys Phe Asn Lys Phe Ile Glu Pro Gln Leu		
450	455	460
Asp Glu Asp Leu Thr Ile Tyr Met Lys Asp Lys Ala Leu Ser Pro Lys		
465	470	475
Lys Ser Asn Trp Asp Thr Val Tyr Pro Ala Ser Asn Leu Leu Tyr Arg		
485	490	495
Thr Asn Ala Ser Asn Glu Ser Arg Arg Leu Val Glu Val Phe Ile Ala		
500	505	510
Asp Ser Lys Phe Asp Pro His Gln Ile Leu Asp Tyr Val Glu Ser Gly		
515	520	525
Asp Trp Leu Asp Asp Pro Glu Phe Asn Ile Ser Tyr Ser Leu Lys Glu		
530	535	540
Lys Glu Ile Lys Gln Glu Gly Arg Leu Phe Ala Lys Met Thr Tyr Lys		
545	550	555
Met Arg Ala Thr Gln Val Leu Ser Glu Thr Leu Leu Ala Asn Asn Ile		
565	570	575
Gly Lys Phe Phe Gln Glu Asn Gly Met Val Lys Gly Glu Ile Glu Leu		
580	585	590
Leu Lys Arg Leu Thr Thr Ile Ser Ile Ser Gly Val Pro Arg Tyr Asn		
595	600	605
Glu Val Tyr Asn Asn Ser Lys Ser His Thr Asp Asp Leu Lys Thr Tyr		
610	615	620

SUBSTITUTE SHEET (RULE 26)



- 249 -

Asn	Lys	Ile	Ser	Asn	Leu	Asn	Leu	Ser	Ser	Asn	Gln	Lys	Ser	Lys	Lys	625	630	635	640
Phe	Glu	Phe	Lys	Ser	Thr	Asp	Ile	Tyr	Asn	Asp	Gly	Tyr	Glu	Thr	Val	645	650	655	
Ser	Cys	Phe	Leu	Thr	Thr	Asp	Leu	Lys	Lys	Tyr	Cys	Leu	Asn	Trp	Arg	660	665	670	
Tyr	Glu	Ser	Thr	Ala	Leu	Phe	Gly	Glu	Thr	Cys	Asn	Gln	Ile	Phe	Gly	675	680	685	
Leu	Asn	Lys	Leu	Phe	Asn	Trp	Leu	His	Pro	Arg	Leu	Glu	Gly	Ser	Thr	690	695	700	
Ile	Tyr	Val	Gly	Asp	Pro	Tyr	Cys	Pro	Pro	Ser	Asp	Lys	Glu	His	Ile	705	710	715	720
Ser	Leu	Glu	Asp	His	Pro	Asp	Ser	Gly	Phe	Tyr	Val	His	Asn	Pro	Arg	725	730	735	
Gly	Gly	Ile	Glu	Gly	Phe	Cys	Gln	Lys	Leu	Trp	Thr	Leu	Ile	Ser	Ile	740	745	750	
Ser	Ala	Ile	His	Leu	Ala	Ala	Val	Arg	Ile	Gly	Val	Arg	Val	Thr	Ala	755	760	765	
Met	Val	Gln	Gly	Asp	Asn	Gln	Ala	Ile	Ala	Val	Thr	Thr	Arg	Val	Pro	770	775	780	
Asn	Asn	Tyr	Asp	Tyr	Arg	Val	Lys	Lys	Glu	Ile	Val	Tyr	Lys	Asp	Val	785	790	795	800
Val	Arg	Phe	Phe	Asp	Ser	Leu	Arg	Glu	Val	Met	Asp	Asp	Leu	Gly	His	805	810	815	
Glu	Leu	Lys	Leu	Asn	Glu	Thr	Ile	Ile	Ser	Ser	Lys	Met	Phe	Ile	Tyr	820	825	830	
Ser	Lys	Arg	Ile	Tyr	Tyr	Asp	Gly	Arg	Ile	Leu	Pro	Gln	Ala	Leu	Lys	835	840	845	
Ala	Leu	Ser	Arg	Cys	Val	Phe	Trp	Ser	Glu	Thr	Val	Ile	Asp	Glu	Thr	850	855	860	
Arg	Ser	Ala	Ser	Ser	Asn	Leu	Ala	Thr	Ser	Phe	Ala	Lys	Ala	Ile	Glu	865	870	875	880
Asn	Gly	Tyr	Ser	Pro	Val	Leu	Gly	Tyr	Ala	Cys	Ser	Ile	Phe	Lys	Asn	885	890	895	

SUBSTITUTE SHEET (RULE 26)

- 250 -

Ile Gln Gln Leu Tyr Ile Ala Leu Gly Met Asn Ile Asn Pro Thr Ile  
 900 905 910  
 Thr Gln Asn Ile Arg Asp Gln Tyr Phe Arg Asn Pro Asn Trp Met Gln  
 915 920 925  
 Tyr Ala Ser Leu Ile Pro Ala Ser Val Gly Gly Phe Asn Tyr Met Ala  
 930 935 940  
 Met Ser Arg Cys Phe Val Arg Asn Ile Gly Asp Pro Ser Val Ala Ala  
 945 950 955 960  
 Leu Ala Asp Ile Lys Arg Phe Ile Lys Ala Asn Leu Leu Asp Arg Ser  
 965 970 975  
 Val Leu Tyr Arg Ile Met Asn Gln Glu Pro Gly Glu Ser Ser Phe Leu  
 980 985 990  
 Asp Trp Ala Ser Asp Pro Tyr Ser Cys Asn Leu Pro Gln Ser Gln Asn  
 995 1000 1005  
 Ile Thr Thr Met Ile Lys Asn Ile Thr Ala Arg Asn Val Leu Gln Asp  
 1010 1015 1020  
 Ser Pro Asn Pro Leu Leu Ser Gly Leu Phe Thr Asn Thr Met Ile Glu  
 1025 1030 1035 1040  
 Glu Asp Glu Glu Leu Ala Glu Phe Leu Met Asp Arg Lys Val Ile Leu  
 1045 1050 1055  
 Pro Arg Val Ala His Asp Ile Leu Asp Asn Ser Leu Thr Gly Ile Arg  
 1060 1065 1070  
 Asn Ala Ile Ala Gly Met Leu Asp Thr Thr Lys Ser Leu Ile Arg Val  
 1075 1080 1085  
 Gly Ile Asn Arg Gly Gly Leu Thr Tyr Ser Leu Leu Arg Lys Ile Ser  
 1090 1095 1100  
 Asn Tyr Asp Leu Val Gln Tyr Glu Thr Leu Ser Arg Thr Leu Arg Leu  
 1105 1110 1115 1120  
 Ile Val Ser Asp Lys Ile Lys Tyr Glu Asp Met Cys Ser Val Asp Leu  
 1125 1130 1135  
 Ala Ile Ala Leu Arg Gln Lys Met Trp Ile His Leu Ser Gly Gly Arg  
 1140 1145 1150  
 Met Ile Ser Gly Leu Glu Thr Pro Asp Pro Leu Glu Leu Leu Ser Gly  
 1155 1160 1165  
 Val Val Ile Thr Gly Ser Glu His Cys Lys Ile Cys Tyr Ser Ser Asp

SUBSTITUTE SHEET (RULE 26)

- 251 -

1170	1175	1180
Gly Thr Asn Pro Tyr Thr Trp Met Tyr Leu Pro Gly Asn Ile Lys Ile 1185	1190	1195 1200
Gly Ser Ala Glu Thr Gly Ile Ser Ser Leu Arg Val Pro Tyr Phe Gly 1205	1210	1215
Ser Val Thr Asp Glu Arg Ser Glu Ala Gln Leu Gly Tyr Ile Lys Asn 1220	1225	1230
Leu Ser Lys Pro Ala Lys Ala Ala Ile Arg Ile Ala Met Ile Tyr Thr 1235	1240	1245
Trp Ala Phe Gly Asn Asp Glu Ile Ser Trp Met Glu Ala Ser Gln Ile 1250	1255	1260
Ala Gln Thr Arg Ala Asn Phe Thr Leu Asp Ser Leu Lys Ile Leu Thr 1265	1270	1275 1280
Pro Val Ala Thr Ser Thr Asn Leu Ser His Arg Leu Lys Asp Thr Ala 1285	1290	1295
Thr Gln Met Lys Phe Ser Ser Thr Ser Leu Ile Arg Val Ser Arg Phe 1300	1305	1310
Ile Thr Met Ser Asn Asp Asn Met Ser Ile Lys Glu Ala Asn Glu Thr 1315	1320	1325
Lys Asp Thr Asn Leu Ile Tyr Gln Gln Ile Met Leu Thr Gly Leu Ser 1330	1335	1340
Val Phe Glu Tyr Leu Phe Arg Leu Lys Glu Thr Thr Gly His Asn Pro 1345	1350	1355 1360
Ile Val Met His Leu His Ile Glu Asp Glu Cys Cys Ile Lys Glu Ser 1365	1370	1375
Phe Asn Asp Glu His Ile Asn Pro Glu Ser Thr Leu Glu Leu Ile Arg 1380	1385	1390
Tyr Pro Glu Ser Asn Glu Phe Ile Tyr Asp Lys Asp Pro Leu Lys Asp 1395	1400	1405
Val Asp Leu Ser Lys Leu Met Val Ile Lys Asp His Ser Tyr Thr Ile 1410	1415	1420
Asp Met Asn Tyr Trp Asp Asp Thr Asp Ile Ile His Ala Ile Ser Ile 1425	1430	1435 1440
Cys Thr Ala Ile Thr Ile Ala Asp Thr Met Ser Gln Leu Asp Arg Asp 1445	1450	1455

SUBSTITUTE SHEET (RULE 26)

- 252 -

Asn Leu Lys Glu Ile Ile Val Ile Ala Asn Asp Asp Asp Ile Asn Ser  
 1460 1465 1470  
 Leu Ile Thr Glu Phe Leu Thr Leu Asp Ile Leu Val Phe Leu Lys Thr  
 1475 1480 1485  
 Phe Gly Gly Leu Leu Val Asn Gln Phe Ala Tyr Thr Leu Tyr Ser Leu  
 1490 1495 1500  
 Lys Ile Glu Gly Arg Asp Leu Ile Trp Asp Tyr Ile Met Arg Thr Leu  
 1505 1510 1515 1520  
 Arg Asp Thr Ser His Ser Ile Leu Lys Val Leu Ser Asn Ala Leu Ser  
 1525 1530 1535  
 His Pro Lys Val Phe Lys Arg Phe Trp Asp Cys Gly Val Leu Asn Pro  
 1540 1545 1550  
 Ile Tyr Gly Pro Asn Thr Ala Ser Gln Asp Gln Ile Lys Leu Ala Leu  
 1555 1560 1565  
 Ser Ile Cys Glu Tyr Ser Leu Asp Leu Phe Met Arg Glu Trp Leu Asn  
 1570 1575 1580  
 Gly Val Ser Leu Glu Ile Tyr Ile Cys Asp Ser Asp Met Glu Val Ala  
 1585 1590 1595 1600  
 Asn Asp Arg Lys Gln Ala Phe Ile Ser Arg His Leu Ser Phe Val Cys  
 1605 1610 1615  
 Cys Leu Ala Glu Ile Ala Ser Phe Gly Pro Asn Leu Leu Asn Leu Thr  
 1620 1625 1630  
 Tyr Leu Glu Arg Leu Asp Leu Leu Lys Gln Tyr Leu Glu Leu Asn Ile  
 1635 1640 1645  
 Lys Glu Asp Pro Thr Leu Lys Tyr Val Gln Ile Ser Gly Leu Leu Ile  
 1650 1655 1660  
 Lys Ser Phe Pro Ser Thr Val Thr Tyr Val Arg Lys Thr Ala Ile Lys  
 1665 1670 1675 1680  
 Tyr Leu Arg Ile Arg Gly Ile Ser Pro Pro Glu Val Ile Asp Asp Trp  
 1685 1690 1695  
 Asp Pro Val Glu Asp Glu Asn Met Leu Asp Asn Ile Val Lys Thr Ile  
 1700 1705 1710  
 Asn Asp Asn Cys Asn Lys Asp Asn Lys Gly Asn Lys Ile Asn Asn Phe  
 1715 1720 1725

SUBSTITUTE SHEET (RULE 26)

- 253 -

Trp Gly Leu Ala Leu Lys Asn Tyr Gln Val Leu Lys Ile Arg Ser Ile  
 1730 1735 1740

Thr Ser Asp Ser Asp Asp Asn Asp Arg Leu Asp Ala Asn Thr Ser Gly  
 1745 1750 1755 1760

Leu Thr Leu Pro Gln Gly Gly Asn Tyr Leu Ser His Gln Leu Arg Leu  
 1765 1770 1775

Phe Gly Ile Asn Ser Thr Ser Cys Leu Lys Ala Leu Glu Leu Ser Gln  
 1780 1785 1790

Ile Leu Met Lys Glu Val Asn Lys Asp Lys Asp Arg Leu Phe Leu Gly  
 1795 1800 1805

Glu Gly Ala Gly Ala Met Leu Ala Cys Tyr Asp Ala Thr Leu Gly Pro  
 1810 1815 1820

Ala Val Asn Tyr Tyr Asn Ser Gly Leu Asn Ile Thr Asp Val Ile Gly  
 1825 1830 1835 1840

Gln Arg Glu Leu Lys Ile Phe Pro Ser Glu Val Ser Leu Val Gly Lys  
 1845 1850 1855

Lys Leu Gly Asn Val Thr Gln Ile Leu Asn Arg Val Lys Val Leu Phe  
 1860 1865 1870

Asn Gly Asn Pro Asn Ser Thr Trp Ile Gly Asn Met Glu Cys Glu Ser  
 1875 1880 1885

Leu Ile Trp Ser Glu Leu Asn Asp Lys Ser Ile Gly Leu Val His Cys  
 1890 1895 1900

Asp Met Glu Gly Ala Ile Gly Lys Ser Glu Glu Thr Val Leu His Glu  
 1905 1910 1915 1920

His Tyr Ser Val Ile Arg Ile Thr Tyr Leu Ile Gly Asp Asp Asp Val  
 1925 1930 1935

Val Leu Val Ser Lys Ile Ile Pro Thr Ile Thr Pro Asn Trp Ser Arg  
 1940 1945 1950

Ile Leu Tyr Leu Tyr Lys Leu Tyr Trp Lys Asp Val Ser Ile Ile Ser  
 1955 1960 1965

Leu Lys Thr Ser Asn Pro Ala Ser Thr Glu Leu Tyr Leu Ile Ser Lys  
 1970 1975 1980

Asp Ala Tyr Cys Thr Ile Met Glu Pro Ser Glu Ile Val Leu Ser Lys  
 1985 1990 1995 2000

Leu Lys Arg Leu Ser Leu Leu Glu Glu Asn Asn Leu Leu Lys Trp Ile

SUBSTITUTE SHEET (RULE 26)

- 254 -

2005	2010	2015
Ile Leu Ser Lys Lys Arg Asn Asn Glu Trp Leu His His Glu Ile Lys 2020 2025 2030		
Glu Gly Glu Arg Asp Tyr Gly Ile Met Arg Pro Tyr His Met Ala Leu 2035 2040 2045		
Gln Ile Phe Gly Phe Gln Ile Asn Leu Asn His Leu Ala Lys Glu Phe 2050 2055 2060		
Leu Ser Thr Pro Asp Leu Thr Asn Ile Asn Asn Ile Ile Gln Ser Phe 2065 2070 2075 2080		
Gln Arg Thr Ile Lys Asp Val Leu Phe Glu Trp Ile Asn Ile Thr His 2085 2090 2095		
Asp Asp Lys Arg His Lys Leu Gly Gly Arg Tyr Asn Ile Phe Pro Leu 2100 2105 2110		
Lys Asn Lys Gly Lys Leu Arg Leu Leu Ser Arg Arg Leu Val Leu Ser 2115 2120 2125		
Trp Ile Ser Leu Ser Leu Ser Thr Arg Leu Leu Thr Gly Arg Phe Pro 2130 2135 2140		
Asp Glu Lys Phe Glu His Arg Ala Gln Thr Gly Tyr Val Ser Leu Ala 2145 2150 2155 2160		
Asp Thr Asp Leu Glu Ser Leu Lys Leu Leu Ser Lys Asn Ile Ile Lys 2165 2170 2175		
Asn Tyr Arg Glu Cys Ile Gly Ser Ile Ser Tyr Trp Phe Leu Thr Lys 2180 2185 2190		
Glu Val Lys Ile Leu Met Lys Leu Ile Gly Gly Ala Lys Leu Leu Gly 2195 2200 2205		
Ile Pro Arg Gln Tyr Lys Glu Pro Glu Asp Gln Leu Leu Glu Asn Tyr 2210 2215 2220		
Asn Gln His Asp Glu Phe Asp Ile Asp 2225 2230		

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15462 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

SUBSTITUTE SHEET (RULE 26)

- 255 -

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACCAAACAAG AGAAGAAACT TGCTTGGTAA TATAAATTTA ACTTAAAATT AACTTAGGAT	60
TTAAGACATT GACTAGAAGG TCAAGAAAAG GGAAGTCTAT AATTTCAAAA ATGTTGAGCC	120
TATTTGATAC ATTTAATGCA CGTAGGCAAG AAAACATAAC AAAATCAGCC GGTGGAGCTA	180
TCATTCCTGG ACAGAAAAAT ACTGTCTCTA TATTCGCCCT TGGACCGACA ATAAGTATG	240
ATAATGAGAA AATGACATTA GCTCTTCTAT TTCTATCTCA TTCACTAGAT AATGAGAAAC	300
AACATGCACA AAGGGCAGGG TTCTTGGTGT CTTTATTGTC AATGGCTTAT GCCAATCCAG	360
AGCTCTACCT AACACAAAT GGAAGTAATG CAGATGCCAA GTATGTCATA TACATGATTG	420
AGAAAGATCT AAAACGGCAA AAGTATGGAG GATTTGTGGT TAAGACGAGA GAGATGATAT	480
ATGAAAAGAC AACTGATTGG ATATTTGGAA GTGACCTGGA TTATGATCAG GAAACTATGT	540
TGCAGAACGG CAGGAACAAT TCAACAATTG AAGACCTTGT CCACACATT TGGTATCCAT	600
CATGTTTAGG AGCTCTTATA ATACAGATCT GGATAGTTCT GGTCAAAGCT ATCACTAGTA	660
TCTCAGGGTT AAGAAAAGGC TTTTTCACCC GATTGGAAGC TTTCAGACAA GATGGAACAG	720
TGCAGGCAGG GCTGGTATTG AGCGGTGACA CAGTGGATCA GATTGGGTCA ATCATGCGGT	780
CTCAACAGAG CTTGGTAACT CTTATGGTTG AAACATTAAT AACAATGAAT ACCAGCAGAA	840
ATGACCTCAC AACCATAGAA AAGAATATAC AAATTGTTGG CAACTACATA AGAGATGCAG	900
GTCTCGCTTC ATTCTTCAAT ACAATCAGAT ATGGAATTGA GACCAGAATG GCAGCTTTGA	960
CTCTATCCAC TCTCAGACCA GATATCAATA GATTAAAAGC TTTGATGGAA CTGTATTTAT	1020
CAAAGGGACC ACGCGCTCCT TTCATCTGTA TCCTCAGAGA TCCTATACAT GGTGAGTTTG	1080
CACCAGGCAA CTATCCTGCC ATATGGAGCT ATGCAATGGG GGTGGCAGTT GTACAAAATA	1140
GAGCCATGCA ACAGTATGTG ACGGGAAGAT CATATCTAGA CATTGATATG TTCCAGCTAG	1200
GACAAGCAGT AGCACGTGAT GCCGAAGCTC AAATGAGCTC AACACTGGAA GATGAACTTG	1260
GAGTGACACA CGAAGCTAAA GAAAGCTTGA AGAGACATAT AAGGAACATA AACAGTTCAG	1320
AGACATCTTT CCACAAACCG ACAGGTGGAT CAGCCATAGA GATGGCAATA GATGAAGAGC	1380

SUBSTITUTE SHEET (RULE 26)

- 256 -

CAGAACAATT CGAACATAGA GCAGATCAAG AACAAAATGG AGAACCTCAA TCATCCATAA	1440
TTCAATATGC CTGGGCAGAA GGAAATAGAA GCGATGATCA GACTGAGCAA GCTACAGAAT	1500
CTGACAATAT CAAGACCGAA CAACAAAACA TCAGAGACAG ACTAAACAAG AGACTCAACG	1560
ACAAGAAGAA ACAAAGCAGT CAACCACCCA CTAATCCCAC AAACAGAACA AACCAGGACG	1620
AAATAGATGA TCTGTTTAAAC GCATTTGGAA GCAACTAATC GAATCAACAT TTTAATCTAA	1680
ATCAATAATA AATAAGAAAA ACTTAGGATT AAAGAATCCT ATCATAACGG AATATAGGGT	1740
GGTAAATTTA GAGTCTGCTT GAAACTCAAT CAATAGAGAG TTGATGGAAA GCGATGCTAA	1800
AAACTATCAA ATCATGGATT CTTGGGAAGA GGAATCAAGA GATAAATCAA CTAATATCTC	1860
CTCGGCCCTC AACATCATTG AATTCATACT CAGCACCGAC CCCCAAGAAG ACTTATCGGA	1920
AAACGACACA ATCAACACAA GAACCCAGCA ACTCAGTGCC ACCATCTGTC AACCAGAAAT	1980
CAAAACCAACA GAAACAAGTG AGAAAGATAG TGGATCAACT GACAAAAATA GACAGTCTGG	2040
GTCATCACAC GAATGTACAA CAGAAGCAAA AGATAGAAAC ATTGATCAGG AAAGTGTACA	2100
GAGAGGACCT GGGAGAAGAA GCAGCTCAGA TAGTAGAGCT GAGACTGTGG TCTCTGGAGG	2160
AATCCCCAGA AGCATCACAG ATTCTAAAAA TGGAACCCAA AACACGGAGG ATATTGATCT	2220
CAATGAAATT AGAAAGATGG ATAAGGACTC TATTGAGGGG AAAATGCGAC AATCTGCAAA	2280
TGTTCCAAGC GAGATATCAG GAAGTGATGA CATATTTACA ACAGAACAAA GTAGAAACAG	2340
TGATCATGGA AGAAGCCTGG AATCTATCAG TACACCTGAT ACAAGATCAA TAAGTGTGTG	2400
TACTGCTGCA ACACCAGATG ATGAAGAAGA AATACTAATG AAAAATAGTA GGACAAAGAA	2460
AAGTTCTTCA ACACATCAAG AAGATGACAA AAGAATTAAA AAAGGGGGAA AAGGGAAAGA	2520
CTGGTTTAAAG AAATCAAAAG ATACCGACAA CCAGATACCA ACATCAGACT ACAGATCCAC	2580
ATCAAAAGGG CAGAAGAAAA TCTCAAAGAC AACCAACCACC AACACCGACA CAAAGGGGCA	2640
AACAGAAATA CAGACAGAAT CATCAGAAAC ACAATCCTCA TCATGGAATC TCATCATCGA	2700
CAACAACACC GACCGGAACG AACAGACAAG CACAACCTCT CCAACAACAA CTTCCAGATC	2760
AACTTATACA AAAGAATCGA TCCGAACAAA CTCTGAATCC AAACCCAAGA CACAAAAGAC	2820
AAATGGAAAG GAAAGGAAGG ATACAGAAGA GAGCAATCGA TTTACAGAGA GGGCAATTAC	2880
TCTATTGCAG AATCTTGGTG TAATTCAATC CACATCAAAA CTAGATTTAT ATCAAGACAA	2940

SUBSTITUTE SHEET (RULE 26)



- 257 -

ACGAGTTGTA TGTGTAGCAA ATGTACTAAA CAATGTAGAT ACTGCATCAA AGATAGATTT	3000
CCTGGCAGGA TTAGTCATAG GGGTTTCAAT GGACAACGAC ACAAATTAA CACAGATACA	3060
AAATGAAATG CTAAACCTCA AAGCAGATCT AAAGAAAATG GACGAATCAC ATAGAAGATT	3120
GATAGAAAAT CAAAGAGAAC AACTGTCATT GATCACGTCA CTAATTTCAA ATCTCAAAAT	3180
TATGACTGAG AGAGGAGGAA AGAAAGACCA AAATGAATCC AATGAGAGAG TATCCATGAT	3240
CAAAACAAAA TTGAAAGAAG AAAAGATCAA GAAGACCAGG TTTGACCCAC TTATGGAGGC	3300
ACAAGGCATT GACAAGAATA TACCCGATCT ATATCGACAT GCAGGAGATA CACTAGAGAA	3360
CGATGTACAA GTTAAATCAG AGATATTAAG TTCATACAAT GAGTCAAATG CAACAAGACT	3420
AATACCCAAA AAAGTGAGCA GTACAATGAG ATCACTAGTT GCAGTCATCA ACAACAGCAA	3480
TCTCTCACA AGCACAAAAC AATCATACAT AAACGAACTC AAACGTTGCA AAAATGATGA	3540
AGAAGTATCT GAATTAATGG ACATGTTCAA TGAAGATGTC AACAAATTGCC AATGATCCAA	3600
CAAAGAAACG ACACCGAACA AACAGACAAG AAACAACAGT AGATCAAAAC CTGTCAACAC	3660
ACACAAAATC AAGCAGAATG AAACAACAGA TATCAATCAA TATACAAATA AGAAAACTT	3720
AGGATTAAAG AATAAATTAA TCCTTGTCCT AAATGAGTAT AACTAACTCT GCAATATACA	3780
CATTCCCAGA ATCATCATT TCTGAAAATG GTCATATAGA ACCATTACCA CTCAAAGTCA	3840
ATGAACAGAG GAAAGCAGTA CCCCACATTA GAGTTGCCAA GATCGGAAAT CCACCAAAC	3900
ACGGATCCCG GTATTTAGAT GTCTTCTTAC TCGGCTTCTT CGAGATGGAA CGAATCAAAG	3960
ACAAATACGG GAGTGTGAAT GATCTCGACA GTGACCCGAG TTACAAAGTT TGTGGCTCTG	4020
GATCATTACC AATCGGATTG GCTAAGTACA CTGGGAATGA CCAGGAATTG TTACAAGCCG	4080
CAACCAAACCT GGATATAGAA GTGAGAAGAA CAGTCAAAGC GAAAGAGATG GTTGTTTACA	4140
CGGTACAAAA TATAAAACCA GAACTGTACC CATGGTCCAA TAGACTAAGA AAAGGAATGC	4200
TGTTGATGTC CAACAAAGTT GCTCTTGCTC CTCAATGTCT TCCACTAGAT AGGAGCATAA	4260
AATTTAGAGT AATCTTCGTG AATTGTACGG CAATTGGATC AATAACCTTG TTCAAATTC	4320
CTAAGTCAAT GGCATCACTA TCTCTAACCA ACACAATATC AATCAATCTG CAGGTACACA	4380
TAAAAACAGG GGTTCCAGACT GATTCTAAAG GGATAGTTCA AATTTTGGAT GAGAAAGGCG	4440
AAAAATCACT GAATTTCATG GTCCATCTCG GATTGATCAA AAGAAAAGTA GGCAGAATGT	4500

SUBSTITUTE SHEET (RULE 26)

- 258 -

ACTCTGTTGA ATACTGTAAA CAGAAAATCG AGAAAATGAG ATTGATATTT TCTTTAGGAC	4560
TAGTTGGAGG AATCAGTCTT CATGTCAATG CAACTGGGTC CATATCAAAA AACTAGCAA	4620
GTCAGCTGGT ATTCAAAAGA GAGATTTGTT ATCCTTTAAT GGATCTAAAT CCGCATCTCA	4680
ATCTAGTTAT CTGGGCTTCA TCAGTAGAGA TTACAAGAGT GGATGCAATT TTCCAACCTT	4740
CTTTACCTGG CGAGTTCAGA TACTATCCTA ATATTATTGC AAAAGGAGTT GGGAAAATCA	4800
AACAATGGAA CTAGTAATCT CTATTTTAGT CCGGACGTAT CTATTAAGCC GAAGCAAATA	4860
AAGGATAATC AAAAAGTTAG GACAAAAGAG GTCAATACCA ACAACTATTA GCAGTCACAC	4920
TCGCAAGAAT AAGAGAGAAG GGACCAAAAA AGTCAAATAG GAGAAATCAA AACAAAAGGT	4980
ACAGAACACC AGAACAACAA AATCAAAACA TCCAATCACC TCAAAACAAA AATTCCAAAA	5040
GAGACCGGCA ACACAACAAG CACTGAACAC AATGCCAACT TCAATACTGC TAATTATTAC	5100
AACCATGATC ATGGCATCTT TCTGCCAAAT AGATATCACA AAACCTACAGC ACGTAGGTGT	5160
ATTGGTCAAC AGTCCCAAAG GGATGAAGAT ATCACAAAAC TTTGAAACAA GATATCTAAT	5220
TTTGAGCCTC ATACCAAAAA TAGAAGACTC TAACTCTTGT GGTGACCAAC AGATCAAGCA	5280
ATACAAGAAG TTATTGGATA GACTGATCAT CCCTTTATAT GATGGATTAA GATTACAGAA	5340
AGATGTGATA GTAACCAATC AAGAATCCAA TGAAAACACT GATCCCAGAA CAAAACGATT	5400
CTTTGGAGGG GTAATTGGAA CCATTGCTCT GGGAGTAGCA ACCTCAGCAC AAATTACAGC	5460
GGCAGTTGCT CTGGTTGAAG CCAAGCAGGC AAGATCAGAC ATCGAAAAAC TCAAAGAAGC	5520
AATTAGGGAC ACAAATAAAG CAGTGCAGTC AGTTCAGAGC TCCATAGGAA ATTTAATAGT	5580
AGCAATTAAA TCAGTCCAGG ATTATGTTAA CAAAGAAATC GTGCCATCGA TTGCGAGGCT	5640
AGGTTGTGAA GCAGCAGGAC TTCAATTAGG AATTGCATTA ACACAGCATT ACTCAGAATT	5700
AACAAACATA TTTGGTGATA ACATAGGATC GTTACAAGAA AAAGGAATAA AATTACAAGG	5760
TATAGCATCA TTATACCGCA CAAATATCAC AGAAATATTC ACAACATCAA CAGTTGATAA	5820
ATATGATATC TATGATCTGT TATTTACAGA ATCAATAAAG GTGAGAGTTA TAGATGTTGA	5880
CTTGAATGAT TACTCAATCA CCCTCCAAGT CAGACTCCCT TTATTAACCTA GGCTGCTGAA	5940
CACTCAGATC TACAAAGTAG ATTCCATATC ATATAACATC CAAAACAGAG AATGGTATAT	6000
CCCTCTTCCC AGCCATATCA TGACGAAAGG GGCATTTCTA GGTGGAGCAG ACGTCAAAGA	6060

SUBSTITUTE SHEET (RULE 26)

- 259 -

ATGTATAGAA GCATTCAGCA GCTATATATG CCCTTCTGAT CCAGGATTGG TATTAAACCA	6120
TGAAATAGAG AGCTGCTTAT CAGGAAACAT ATCCCAATGT CCAAGAACAA CGGTCACATC	6180
AGACATTGTT CCAAGATATG CATTGTGCAA TGGAGGAGTG GTTGCAAACGT GTATAACAAC	6240
CACCTGTACA TGCAACGGAA TTGGTAATAG AATCAATCAA CCACCTGATC AAGGAGTAAA	6300
AATTATAACA CATAAAGAAT GTAGTACAGT AGGTATCAAC GGAATGCTGT TCAATACAAA	6360
TAAAGAAGGA ACTCTTGCAAT TCTATACACC AAATGATATA AACTAAACA ATTCTGTTAC	6420
ACTTGATCCA ATTGACATAT CAATCGAGCT CAACAAGGCC AAATCAGATC TAGAAGAATC	6480
AAAAGAATGG ATAAGAAGGT CAAATCAAAA ACTAGATTCT ATTGGAAATT GGCATCAATC	6540
TAGCACTACA ATCATAATTA TTTTGATAAT GATCATTATA TTGTTTATAA TTAATATAAC	6600
GATAATTACA ATTGCAATTA AGTATTACAG AATTCAAAAG AGAAATCGAG TGGATCAAAA	6660
TGACAAGCCA TATGTACTAA CAAACAAATA ACATATCTAC AGATCATTAG ATATTAAAAT	6720
TATAAAAAAC TTAGGAGTAA AGTTACGCAA TCCAACCTCTA CTCATATAAT TGAGGAAGGA	6780
CCCAATAGAC AAATCCAAAT TCGAGATGGA ATACTGGAAG CATACCAATC ACGGAAAGGA	6840
TGCTGGCAAT GAGCTGGAGA CGTCTATGGC TACTCATGGC AACAAAGCTCA CTAATAAGAT	6900
AATATACATA TTATGGACAA TAATCCTGGT GTTATTATCA ATAGTCTTCA TCATAGTGCT	6960
AATTAATTCC ATCAAAAGTG AAAAGGCCCA CGAATCATTG CTGCAAGACA TAAATAATGA	7020
GTTTATGGAA ATTACAGAAA AGATCCAAAT GGCATCGGAT AATACCAATG ATCTAATACA	7080
GTCAGGAGTG AATACAAGGC TTCTTACAAT TCAGAGTCAT GTCCAGAATT ACATACCAAT	7140
ATCATTGACA CAACAGATGT CAGATCTTAG GAAATTCATT AGTGAAATTA CAATTAGAAA	7200
TGATAATCAA GAAGTGCTGC CACAAAGAAT AACACATGAT GTAGGTATAA AACCTTTAAA	7260
TCCAGATGAT TTTTGGAGAT GCACGTCTGG TCTTCCATCT TTAATGAAAA CTCCAAAAAT	7320
AAGGTAAATG CCAGGGCCGG GATTATTAGC TATGCCAAGC ACTGTTGATG GCTGTGTTAG	7380
AACTCCGTCT TTAGTTATAA ATGATCTGAT TTATGCTTAT ACCTCAAATC TAATTACTCG	7440
AGGTTGTCAG GATATAGGAA AATCATATCA AGTCTTACAG ATAGGGATAA TAACTGTAAA	7500
CTCAGACTTG GTACCTGACT TAAATCCTAG GATCTCTCAT ACCTTTAACA TAAATGACAA	7560
TAGGAAGTCA TGTCTCTAG CACTCCTAAA TACAGATGTA TATCAACTGT GTTCAACTCC	7620

SUBSTITUTE SHEET (RULE 26)

- 260 -

CAAAGTTGAT GAAAGATCAG ATTATGCATC ATCAGGCATA GAAGATATTG TACTTGATAT	7680
TGTCAATTAT GATGGTTCAA TCTCAACAAC AAGATTTAAG AATAATAACA TAAGCTTTGA	7740
TCAACCATAT GCTGCACTAT ACCCATCTGT TGGACCAGGG ATATACTACA AAGGCAAAAT	7800
AATATTTCTC GGGTATGGAG GTCTTGAACA TCCAATAAAT GAGAATGTAA TCTGCAACAC	7860
AACTGGGTGC CCCGGGAAAA CACAGAGAGA CTGTAATCAA GCGTCTCATA GTCCATGGTT	7920
TTCAGATAGG AGGATGGTCA ACTCCATCAT TGTGCTGAC AAAGGCTTAA ACTCAATTCC	7980
AAAATTGAAA GTATGGACGA TATCTATGCG ACAAATTAC TGGGGGTCAG AAGGAAGGTT	8040
ACTTCTACTA GGTAACAAGA TCTATATATA TACAAGATCT ACAAGTTGGC ATAGCAAGTT	8100
ACAATTAGGA ATAATTGATA TTAGTGATTA CAGTGATATA AGGATAAAAT GGACATGGCA	8160
TAATGTGCTA TCAAGACCAG GAAACAATGA ATGTCCATGG GGACATTCAT GTCCAGATGG	8220
ATGTATAACA GGAGTATATA CTGATGCATA TCCACTCAAT CCCACAGGGA GCATTGTGTC	8280
ATCTGTCATA TTAGACTCAC AAAAATCGAG AGTGAACCCA GTCATAACTT ACTCAACAGC	8340
AACCGAAAGA GTAAACGAGC TGGCCATCCT AAACAGAACA CTCTCAGCTG GATATACAAC	8400
AACAAGCTGC ATTACACACT ATAACAAAGG ATATTGTTTT CATATAGTAG AAATAAATCA	8460
TAAAAGCTTA AACACATTTT AACCCATGTT GTTCAAAACA GAGATTCCAA AAAGCTGCAG	8520
TTAATCATAA TTAACCATAA TATGCATCAA TCTATCTATA ATACAAGTAT ATGATAAGTA	8580
ATCAGCAATC AGACAATAGA CAAAAGGGAA ATATAAAAAA CTTAGGAGCA AAGCGTGCTC	8640
GGGAAATGGA CACTGAATCT AACAAATGGCA CTGTATCTGA CATACTCTAT CCTGAGTGTC	8700
ACCTTAACTC TCCTATCGTT AAAGGTAAAA TAGCACAATT ACACACTATT ATGAGTCTAC	8760
CTCAGCCTTA TGATATGGAT GACGACTCAA TACTAGTTAT CACTAGACAG AAAATAAAAC	8820
TTAATAAATT GGATAAAAGA CAACGATCTA TTAGAAGATT AAAATTAATA TTAAGTAAA	8880
AAGTGAATGA CTTAGGAAAA TACACATTTA TCAGATATCC AGAAATGTCA AAAGAAATGT	8940
TCAAATTATA TATACCTGGT ATTAACAGTA AAGTGAATGA ATTATTACTT AAAGCAGATA	9000
GAACATATAG TCAAATGACT GATGGATTAA GAGATCTATG GATTAATGTG CTATCAAAAT	9060
TAGCCTCAAA AAATGATGGA AGCAATTATG ATCTTAATGA AGAAATTAAT AATATATCGA	9120
AAGTTCACAC AACCTATAAA TCAGATAAAT GGTATAATCC ATTCAAAACA TGGTTTACTA	9180

SUBSTITUTE SHEET (RULE 26)

- 261 -

TCAAGTATGA TATGAGAAGA TTACAAAAAG CTCGAAATGA GATCACTTTT AATGTTGGGA	9240
AGGATTATAA CTTGTTAGAA GACCAGAAGA ATTTCTTATT GATACATCCA GAATTGGTTT	9300
TGATATTAGA TAAACAAAAC TACAATGGTT ATCTAATTAC TCCTGAATTA GTATTGATGT	9360
ATTGTGACGT AGTCGAAGGC CGATGGAATA TAAGTGCATG TGCTAAGTTA GATCCAAAAT	9420
TACAATCTAT GTATCAGAAA GGTAATAACC TGTGGGAAGT GATAGATAAA TTGTTTCCAA	9480
TTATGGGAGA AAAGACATTT GATGTGATAT CGTTATTAGA ACCACTTGCA TTATCCTTAA	9540
TTCAAACCTCA TGATCCTGTT AAACAACATA GAGGAGCTTT TTTAAATCAT GTGTTATCCG	9600
AGATGGAATT AATATTTGAA TCTAGAGAAT CGATTAAGGA ATTTCTGAGT GTAGATTACA	9660
TTGATAAAAT TTTAGATATA TTTAATAAGT CTACAATAGA TGAAATAGCA GAGATTTTCT	9720
CTTTTTTTAG AACATTTGGG CATCCTCCAT TAGAAGCTAG TATTGCAGCA GAAAAGGTTA	9780
GAAAATATAT GTATATTGGA AAACAATTAA AATTTGACAC TATTAATAAA TGTCATGCTA	9840
TCTTCTGTAC AATAATAATT AACGGATATA GAGAGAGGCA TGGTGGACAG TGGCCTCCTG	9900
TGACATTACC TGATCATGCA CACGAATTCA TCATAAATGC TTACGGTTCA AACTCTGCGA	9960
TATCATATGA GAATGCTGTT GATTATTACC AGAGCTTTAT AGGAATAAAA TTCAATAAAT	10020
TCATAGAGCC TCAGTTAGAT GAGGATTTGA CAATTTATAT GAAAGATAAA GCATTATCTC	10080
CAAAAAAATC AAATTGGGAC ACAGTTTATC CTGCATCTAA TTTACTGTAC CGTACTAACG	10140
CATCCAACGA ATCACGAAGA TTAGTTGAAG TATTTATAGC AGATAGTAAA TTTGATCCTC	10200
ATCAGATATT GGATTATGTA GAATCTGGGG ACTGGTTAGA TGATCCAGAA TTTAATATT	10260
CTTATAGTCT TAAAGAAAAA GAGATCAAAC AGGAAGGTAG ACTCTTTGCA AAAATGACAT	10320
ACAAAATGAG AGCTACACAA GTTTTATCAG AGACACTACT TGCAAATAAC ATAGGAAAAT	10380
TCTTTCAAGA AAATGGGATG GTGAAGGGAG AGATTGAATT ACTTAAGAGA TTAACAACCA	10440
TATCAATATC AGGAGTTCCA CGGTATAATG AAGTGACAA TAATTCTAAA AGCCATACAG	10500
ATGACCTTAA AACCTACAAT AAAATAAGTA ATCTTAATTT GTCTTCTAAT CAGAAATCAA	10560
AGAAATTTGA ATTCAAGTCA ACGGATATCT ACAATGATGG ATACGAGACT GTGAGCTGTT	10620
TCCTAACAA ACAGTCTCAA AAATACTGTC TTAATTGGAG ATATGAATCA ACAGCTCTAT	10680
TTGGAGAAAC TTGCAACCAA ATATTTGGAT TAAATAAATT GTTTAATTGG TTACACCCTC	10740

SUBSTITUTE SHEET (RULE 26)

- 262 -

GTCTTGAAGG AAGTACAATC TATGTAGGTG ATCCTTACTG TCCTCCATCA GATAAAGAAC	10800
ATATATCATT AGAGGATCAC CCTGATTCTG GTTTTTACGT TCATAACCCA AGAGGGGGTA	10860
TAGAAGGATT TTGTCAAAAA TTATGGACAC TCATATCTAT AAGTGCAATA CATCTAGCAG	10920
CTGTTAGAAT AGGCGTGAGG GTGACTGCAA TGGTTCAAGG AGACAATCAA GCTATAGCTG	10980
TAACCACAAG AGTACCCAAC AATTATGACT ACAGAGTTAA GAAGGAGATA GTTTATAAAG	11040
ATGTAGTGAG ATTTTTTGAT TCATTAAGAG AAGTGATGGA TGATCTAGGT CATGAACTTA	11100
AATTAATGA AACGATTATA AGTAGCAAGA TGTTCATATA TAGCAAAAGA ATCTATTATG	11160
ATGGGAGAAT TCTTCCTCAA GCTCTAAAAG CATTATCTAG ATGTGTCTTC TGGTCAGAGA	11220
CAGTAATAGA CGAAACAAGA TCAGCATCTT CAAATTTGGC AACATCATTT GCAAAAGCAA	11280
TTGAGAATGG TTATTCACCT GTTCTAGGAT ATGCATGCTC AATTTTTAAG AACATTCAAC	11340
AACTATATAT TGCCCTTGGG ATGAATATCA ATCCAACAT AACACAGAAT ATCAGAGATC	11400
AGTATTTTAG GAATCCAAAT TGGATGCAAT ATGCCTCTTT AATACCTGCT AGTGTGGGG	11460
GATTCAATCA CATGGCCATG TCAAGATGTT TTGTAAGGAA TATTGGTGAT CCATCAGTTG	11520
CCGCATTGGC TGATATTAAA AGATTATTA AGGCGAATCT ATTAGACCGA AGTGTTCTTT	11580
ATAGGATTAT GAATCAAGAA CCAGGTGAGT CATCTTTTTT TGAAGGGCT TCAGATCCAT	11640
ATTCATGCAA TTTACCACAA TCTCAAATA TAACCACCAT GATAAAAAAT ATAACAGCAA	11700
GGAATGTATT ACAAGATTCA CCAAATCCAT TATTATCTGG ATTATTCACA AATACAATGA	11760
TAGAAGAAGA TGAAGAATTA GCTGAGTTCC TGATGGACAG GAAGGTAATT CTCCCTAGAG	11820
TTGCACATGA TATTCTAGAT AATTCTCTCA CAGGAATTAG AAATGCCATA GCTGGAATGT	11880
TAGATACGAC AAAATCACTA ATTCGGGTTG GCATAAATAG AGGAGGACTG ACATATAGTT	11940
TGTTGAGGAA AATCAGTAAT TACGATCTAG TACAATATGA AACACTAAGT AGGACTTTGC	12000
GACTAATTGT AAGTGATAAA ATCAAGTATG AAGATATGTG TTCGGTAGAC CTTGCCATAG	12060
CATTGCGACA AAAGATGTGG ATTCATTTAT CAGGAGGAAG GATGATAAGT GGACTTGAAA	12120
CGCCTGACCC ATTAGAATTA CTATCTGGGG TAGTAATAAC AGGATCAGAA CATTGTAAAA	12180
TATGTTATTC TTCAGATGGC ACAAACCCAT ATACTTGGAT GTATTTACCC GGTAATATCA	12240
AAATAGGATC AGCAGAAACA GGTATATCGT CATTAGAGT TCCTTATTTT GGATCAGTCA	12300

SUBSTITUTE SHEET (RULE 26)

- 263 -

CTGATGAAAG ATCTGAAGCA CAATTAGGAT ATATCAAGAA TCTTAGTAAA CCTGCAAAAG	12360
CCGCAATAAG AATAGCAATG ATATATACAT GGGCATTGAG TAATGATGAG ATATCTTGGA	12420
TGGAAGCCTC ACAGATAGCA CAAACACGTG CAAATTTTAC ACTAGATAGT CTCAAAATTT	12480
TAACACCGGT AGCTACATCA ACAAATTTAT CACACAGATT AAAGGATACT GCAACTCAGA	12540
TGAAATTCTC CAGTACATCA TTGATCAGAG TCAGCAGATT TATAACAATG TCCAATGATA	12600
ACATGTCTAT CAAAGAAGCT AATGAAACCA AAGATACTAA TCTTATTTAT CAACAAATAA	12660
TGTTAACAGG ATTAAGTGTT TTCGAATATT TATTTAGATT AAAAGAAACC ACAGGACACA	12720
ACCCTATAGT TATGCATCTG CACATAGAAG ATGAGTGTG TATTAAAGAA AGTTTTAATG	12780
ATGAACATAT TAATCCAGAG TCTACATTAG AATTAATTCG ATATCCTGAA AGTAATGAAT	12840
TTATTTATGA TAAAGACCCA CTCAAAGATG TGGACTTATC AAAACTTATG GTTATTAAAG	12900
ACCATTCTTA CACAATTGAT ATGAATTATT GGGATGATAC TGACATCATA CATGCAATTT	12960
CAATATGTAC TGCAATTACA ATAGCAGATA CTATGTCACA ATTAGATCGA GATAATTTAA	13020
AAGAGATAAT AGTTATTGCA AATGATGATG ATATTAATAG CTTAATCACT GAATTTTGA	13080
CTCTTGACAT ACTTGATTTT CTCAAGACAT TTGGTGGATT ATTAGTAAAT CAATTTGCAT	13140
ACACTCTTTA TAGTCTAAAA ATAGAAGGTA GGGATCTCAT TTGGGATTAT ATAATGAGAA	13200
CACTGAGAGA TACTTCCCAT TCAATATTAA AAGTATTATC TAATGCATTA TCTCATCCTA	13260
AAGTATTCAA GAGGTTCTGG GATTGTGGAG TTTTAAACCC TATTTATGGT CCTAATATTG	13320
CTAGTCAAGA CCAGATAAAA CTTGCCCTAT CTATATGTGA ATATTCACCTA GATCTATTTA	13380
TGAGAGAATG GTTGAATGGT GTATCACTTG AAATATACAT TTGTGACAGC GATATGGAAG	13440
TTGCAAATGA TAGGAAACAA GCCTTTATTT CTAGACACCT TTCATTGTT TGTGTTTAG	13500
CAGAAATTGC ATCTTTCGGA CCTAACCTGT TAACTTAAC ATACTTGAG AGACTTGATC	13560
TATTGAAACA ATATCTTGAA TTAAATATTA AAGAAGACCC TACTCTTAAA TATGTACAAA	13620
TATCTGGATT ATTAATTAAA TCGTTCCCAT CAACTGTAAC ATACGTAAGA AAGACTGCAA	13680
TCAAATATCT AAGGATTGCG GGTATTAGTC CACCTGAGGT AATTGATGAT TGGGATCCGG	13740
TAGAAGATGA AAATATGCTG GATAACATTG TCAAACTAT AAATGATAAC TGTAATAAAG	13800
ATAATAAAGG GAATAAAATT AACAATTTCT GGGGACTAGC ACTTAAGAAC TATCAAGTCC	13860

SUBSTITUTE SHEET (RULE 26)

- 264 -

TTAAAATCAG ATCTATAACA AGTGATTCTG ATGATAATGA TAGACTAGAT GCTAATACAA	13920
GTGGTTTGAC ACTTCCTCAA GGAGGGAATT ATCTATCGCA TCAATTGAGA TTATTCGGAA	13980
TCAACAGCAC TAGTTGTCTG AAAGCTCTTG AGTTATCACA AATTTTAATG AAGGAAGTCA	14040
ATAAAGACAA GGACAGGCTC TTCCTGGGAG AAGGAGCAGG AGCTATGCTA GCATGTTATG	14100
ATGCCACATT AGGACCTGCA GTTAATTATT ATAATTCAGG TTTGAATATA ACAGATGTAA	14160
TTGGTCAACG AGAATTGAAA ATATTTCTCT CAGAGGTATC ATTAGTAGGT AAAAAATTAG	14220
GAAATGTGAC ACAGATTCTT AACAGGGTAA AAGTACTGTT CAATGGGAAT CCTAATTCAA	14280
CATGGATAGG AAATATGGAA TGTGAGAGCT TAATATGGAG TGAATTAAAT GATAAGTCCA	14340
TTGGATTAGT ACATTGTGAT ATGGAAGGAG CTATCGGTAA ATCAGAAGAA ACTGTTCTAC	14400
ATGAACATTA TAGTGTATA AGAATTACAT ACTTGATTGG GGATGATGAT GTTGTTTTAG	14460
TTTCCAAAAT TATACCTACA ATCACTCCGA ATTGGTCTAG AATACTTTAT CTATATAAAT	14520
TATATTGGAA AGATGTAAGT ATAATATCAC TCAAACTTC TAATCCTGCA TCAACAGAAT	14580
TATATCTAAT TTCGAAAGAT GCATATTGTA CTATAATGGA ACCTAGTGAA ATTGTTTAT	14640
CAAACTTAA AAGATTGTCA CTCTTGAAG AAAATAATCT ATTAATATGG ATCATTTTAT	14700
CAAAGAAGAG GAATAATGAA TGGTTACATC ATGAAATCAA AGAAGGAGAA AGAGATTATG	14760
GAATCATGAG ACCATATCAT ATGGCACTAC AAATCTTTGG ATTTCAAATC AATTAAATC	14820
ATCTGGCGAA AGAATTTTAA TCAACCCAG ATCTGACTAA TATCAACAAT ATAATCCAAA	14880
GTTTTAGCG AACATAAAG GATGTTTTAT TTGAATGGAT TAATATACT CATGATGATA	14940
AGAGACATAA ATTAGGCGGA AGATATAACA TATCCCACT GAAAAATAAG GGAAAGTTAA	15000
GACTGCTATC GAGAAGACTA GTATTAAGTT GGATTCATT ATCATTATCG ACTCGATTAC	15060
TTACAGGTCG CTTTCCTGAT GAAAAATTTG AACATAGAGC ACAGACTGGA TATGTATCAT	15120
TAGCTGATAC TGATTAGAA TCATTAAAGT TATTGTGAA AAACATCATT AAGAATTACA	15180
GAGAGTGAT AGGATCAATA TCATATTGGT TTCTAACCAA AGAAGTTAAA ATACTTATGA	15240
AATTGATTGG TGGTGCTAAA TTATTAGGAA TTCCAGACA ATATAAGAA CCCGAAGACC	15300
AGTTATTAGA AACTACAAT CAACATGATG AATTGATAT CGATTAAAC ATAAATACAA	15360
TGAAGATATA TCCTAACCTT TATCTTTAAG CCTAGGAATA GACAAAAAGT AAGAAAAACA	15420

SUBSTITUTE SHEET (RULE 26)



- 265 -

TGTAATATAT ATATACAAA CAGAGTTCTT CTCTTGTTTGT GT

15462

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Asp Thr Glu Ser Asn Asn Gly Thr Val Ser Asp Ile Leu Tyr Pro
1           5           10           15

Glu Cys His Leu Asn Ser Pro Ile Val Lys Gly Lys Ile Ala Gln Leu
20          25          30

His Thr Ile Met Ser Leu Pro Gln Pro Tyr Asp Met Asp Asp Asp Ser
35          40          45

Ile Leu Val Ile Thr Arg Gln Lys Ile Lys Leu Asn Lys Leu Asp Lys
50          55          60

Arg Gln Arg Ser Ile Arg Arg Leu Lys Leu Ile Leu Thr Glu Lys Val
65          70          75          80

Asn Asp Leu Gly Lys Tyr Thr Phe Ile Arg Tyr Pro Glu Met Ser Lys
85          90          95

Glu Met Phe Lys Leu Tyr Ile Pro Gly Ile Asn Ser Lys Val Thr Glu
100         105         110

Leu Leu Leu Lys Ala Asp Arg Thr Tyr Ser Gln Met Thr Asp Gly Leu
115         120         125

Arg Asp Leu Trp Ile Asn Val Leu Ser Lys Leu Ala Ser Lys Asn Asp
130         135         140

Gly Ser Asn Tyr Asp Leu Asn Glu Glu Ile Asn Asn Ile Ser Lys Val
145         150         155         160

His Thr Thr Tyr Lys Ser Asp Lys Trp Tyr Asn Pro Phe Lys Thr Trp
165         170         175

Phe Thr Ile Lys Tyr Asp Met Arg Arg Leu Gln Lys Ala Arg Asn Glu

```

SUBSTITUTE SHEET (RULE 26)

180

185

190

Ile Thr Phe Asn Val Gly Lys Asp Tyr Asn Leu Leu Glu Asp Gln Lys  
195 200 205

Asn Phe Leu Leu Ile His Pro Glu Leu Val Leu Ile Leu Asp Lys Gln  
210 215 220

Asn Tyr Asn Gly Tyr Leu Ile Thr Pro Glu Leu Val Leu Met Tyr Cys  
225 230 235 240

Asp Val Val Glu Gly Arg Trp Asn Ile Ser Ala Cys Ala Lys Leu Asp  
245 250 255

Pro Lys Leu Gln Ser Met Tyr Gln Lys Gly Asn Asn Leu Trp Glu Val  
260 265 270

Ile Asp Lys Leu Phe Pro Ile Met Gly Glu Lys Thr Phe Asp Val Ile  
275 280 285

Ser Leu Leu Glu Pro Leu Ala Leu Ser Leu Ile Gln Thr His Asp Pro  
290 295 300

Val	Lys	Gln	Leu	Arg	Gly	Ala	Phe	Leu	Asn	His	Val	Leu	Ser	Glu	Met
305					310					315					320

Glu Leu Ile Phe Glu Ser Arg Glu Ser Ile Lys Glu Phe Leu Ser Val  
325 330 335

Asp Tyr Ile Asp Lys Ile Leu Asp Ile Phe Asn Lys Ser Thr Ile Asp  
340 345 350

Glu Ile Ala Glu Ile Phe Ser Phe Phe Arg Thr Phe Gly His Pro Pro  
355 360 365

Leu Glu Ala Ser Ile Ala Ala Glu Lys Val Arg Lys Tyr Met Tyr Ile  
370 375 380

Gly Lys Gln Leu Lys Phe Asp Thr Ile Asn Lys Cys His Ala Ile Phe  
385 390 395 400

Cys Thr Ile Ile Ile Asn Gly Tyr Arg Glu Arg His Gly Gly Gln Trp  
405 410 415

Pro Pro Val Thr Leu Pro Asp His Ala His Glu Phe Ile Ile Asn Ala  
420 425 430

Tyr Gly Ser Asn Ser Ala Ile Ser Tyr Glu Asn Ala Val Asp Tyr Tyr  
435 440 445

Gln Ser Phe Ile Gly Ile Lys Phe Asn Lys Phe Ile Glu Pro Gln Leu  
450 455 460

BNSDOCID: &lt;WO 9813501A2&gt;

- 267 -

Asp Glu Asp Leu Thr Ile Tyr Met Lys Asp Lys Ala Leu Ser Pro Lys  
 465 470 475 480  
 Lys Ser Asn Trp Asp Thr Val Tyr Pro Ala Ser Asn Leu Leu Tyr Arg  
 485 490 495  
 Thr Asn Ala Ser Asn Glu Ser Arg Arg Leu Val Glu Val Phe Ile Ala  
 500 505 510  
 Asp Ser Lys Phe Asp Pro His Gln Ile Leu Asp Tyr Val Glu Ser Gly  
 515 520 525  
 Asp Trp Leu Asp Asp Pro Glu Phe Asn Ile Ser Tyr Ser Leu Lys Glu  
 530 535 540  
 Lys Glu Ile Lys Gln Glu Gly Arg Leu Phe Ala Lys Met Thr Tyr Lys  
 545 550 555 560  
 Met Arg Ala Thr Gln Val Leu Ser Glu Thr Leu Leu Ala Asn Asn Ile  
 565 570 575  
 Gly Lys Phe Phe Gln Glu Asn Gly Met Val Lys Gly Glu Ile Glu Leu  
 580 585 590  
 Leu Lys Arg Leu Thr Thr Ile Ser Ile Ser Gly Val Pro Arg Tyr Asn  
 595 600 605  
 Glu Val Tyr Asn Asn Ser Lys Ser His Thr Asp Asp Leu Lys Thr Tyr  
 610 615 620  
 Asn Lys Ile Ser Asn Leu Asn Leu Ser Ser Asn Gln Lys Ser Lys Lys  
 625 630 635 640  
 Phe Glu Phe Lys Ser Thr Asp Ile Tyr Asn Asp Gly Tyr Glu Thr Val  
 645 650 655  
 Ser Cys Phe Leu Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670  
 Tyr Glu Ser Thr Ala Leu Phe Gly Glu Thr Cys Asn Gln Ile Phe Gly  
 675 680 685  
 Leu Asn Lys Leu Phe Asn Trp Leu His Pro Arg Leu Glu Gly Ser Thr  
 690 695 700  
 Ile Tyr Val Gly Asp Pro Tyr Cys Pro Pro Ser Asp Lys Glu His Ile  
 705 710 715 720  
 Ser Leu Glu Asp His Pro Asp Ser Gly Phe Tyr Val His Asn Pro Arg  
 725 730 735

SUBSTITUTE SHEET (RULE 26)

- 268 -

Gly Gly Ile Glu Gly Phe Cys Gln Lys Leu Trp Thr Leu Ile Ser Ile  
 740 745 750  
 Ser Ala Ile His Leu Ala Ala Val Arg Ile Gly Val Arg Val Thr Ala  
 755 760 765  
 Met Val Gln Gly Asp Asn Gln Ala Ile Ala Val Thr Thr Arg Val Pro  
 770 775 780  
 Asn Asn Tyr Asp Tyr Arg Val Lys Lys Glu Ile Val Tyr Lys Asp Val  
 785 790 795 800  
 Val Arg Phe Phe Asp Ser Leu Arg Glu Val Met Asp Asp Leu Gly His  
 805 810 815  
 Glu Leu Lys Leu Asn Glu Thr Ile Ile Ser Ser Lys Met Phe Ile Tyr  
 820 825 830  
 Ser Lys Arg Ile Tyr Tyr Asp Gly Arg Ile Leu Pro Gln Ala Leu Lys  
 835 840 845  
 Ala Leu Ser Arg Cys Val Phe Trp Ser Glu Thr Val Ile Asp Glu Thr  
 850 855 860  
 Arg Ser Ala Ser Ser Asn Leu Ala Thr Ser Phe Ala Lys Ala Ile Glu  
 865 870 875 880  
 Asn Gly Tyr Ser Pro Val Leu Gly Tyr Ala Cys Ser Ile Phe Lys Asn  
 885 890 895  
 Ile Gln Gln Leu Tyr Ile Ala Leu Gly Met Asn Ile Asn Pro Thr Ile  
 900 905 910  
 Thr Gln Asn Ile Arg Asp Gln Tyr Phe Arg Asn Pro Asn Trp Met Gln  
 915 920 925  
 Tyr Ala Ser Leu Ile Pro Ala Ser Val Gly Gly Phe Asn His Met Ala  
 930 935 940  
 Met Ser Arg Cys Phe Val Arg Asn Ile Gly Asp Pro Ser Val Ala Ala  
 945 950 955 960  
 Leu Ala Asp Ile Lys Arg Phe Ile Lys Ala Asn Leu Leu Asp Arg Ser  
 965 970 975  
 Val Leu Tyr Arg Ile Met Asn Gln Glu Pro Gly Glu Ser Ser Phe Phe  
 980 985 990  
 Asp Trp Ala Ser Asp Pro Tyr Ser Cys Asn Leu Pro Gln Ser Gln Asn  
 995 1000 1005  
 Ile Thr Thr Met Ile Lys Asn Ile Thr Ala Arg Asn Val Leu Gln Asp

SUBSTITUTE SHEET (RULE 26)

- 269 -

1010	1015	1020
Ser Pro Asn Pro Leu Leu Ser Gly Leu Phe Thr Asn Thr Met Ile Glu 1025	1030	1035 1040
Glu Asp Glu Glu Leu Ala Glu Phe Leu Met Asp Arg Lys Val Ile Leu 1045	1050	1055
Pro Arg Val Ala His Asp Ile Leu Asp Asn Ser Leu Thr Gly Ile Arg 1060	1065	1070
Asn Ala Ile Ala Gly Met Leu Asp Thr Thr Lys Ser Leu Ile Arg Val 1075	1080	1085
Gly Ile Asn Arg Gly Gly Leu Thr Tyr Ser Leu Leu Arg Lys Ile Ser 1090	1095	1100
Asn Tyr Asp Leu Val Gln Tyr Glu Thr Leu Ser Arg Thr Leu Arg Leu 1105	1110	1115 1120
Ile Val Ser Asp Lys Ile Lys Tyr Glu Asp Met Cys Ser Val Asp Leu 1125	1130	1135
Ala Ile Ala Leu Arg Gln Lys Met Trp Ile His Leu Ser Gly Gly Arg 1140	1145	1150
Met Ile Ser Gly Leu Glu Thr Pro Asp Pro Leu Glu Leu Leu Ser Gly 1155	1160	1165
Val Val Ile Thr Gly Ser Glu His Cys Lys Ile Cys Tyr Ser Ser Asp 1170	1175	1180
Gly Thr Asn Pro Tyr Thr Trp Met Tyr Leu Pro Gly Asn Ile Lys Ile 1185	1190	1195 1200
Gly Ser Ala Glu Thr Gly Ile Ser Ser Leu Arg Val Pro Tyr Phe Gly 1205	1210	1215
Ser Val Thr Asp Glu Arg Ser Glu Ala Gln Leu Gly Tyr Ile Lys Asn 1220	1225	1230
Leu Ser Lys Pro Ala Lys Ala Ala Ile Arg Ile Ala Met Ile Tyr Thr 1235	1240	1245
Trp Ala Phe Gly Asn Asp Glu Ile Ser Trp Met Glu Ala Ser Gln Ile 1250	1255	1260
Ala Gln Thr Arg Ala Asn Phe Thr Leu Asp Ser Leu Lys Ile Leu Thr 1265	1270	1275 1280
Pro Val Ala Thr Ser Thr Asn Leu Ser His Arg Leu Lys Asp Thr Ala 1285	1290	1295

SUBSTITUTE SHEET (RULE 26)

- 270 -

Thr Gln Met Lys Phe Ser Ser Thr Ser Leu Ile Arg Val Ser Arg Phe  
 1300 1305 1310  
 Ile Thr Met Ser Asn Asp Asn Met Ser Ile Lys Glu Ala Asn Glu Thr  
 1315 1320 1325  
 Lys Asp Thr Asn Leu Ile Tyr Gln Gln Ile Met Leu Thr Gly Leu Ser  
 1330 1335 1340  
 Val Phe Glu Tyr Leu Phe Arg Leu Lys Glu Thr Thr Gly His Asn Pro  
 1345 1350 1355 1360  
 Ile Val Met His Leu His Ile Glu Asp Glu Cys Cys Ile Lys Glu Ser  
 1365 1370 1375  
 Phe Asn Asp Glu His Ile Asn Pro Glu Ser Thr Leu Glu Leu Ile Arg  
 1380 1385 1390  
 Tyr Pro Glu Ser Asn Glu Phe Ile Tyr Asp Lys Asp Pro Leu Lys Asp  
 1395 1400 1405  
 Val Asp Leu Ser Lys Leu Met Val Ile Lys Asp His Ser Tyr Thr Ile  
 1410 1415 1420  
 Asp Met Asn Tyr Trp Asp Asp Thr Asp Ile Ile His Ala Ile Ser Ile  
 1425 1430 1435 1440  
 Cys Thr Ala Ile Thr Ile Ala Asp Thr Met Ser Gln Leu Asp Arg Asp  
 1445 1450 1455  
 Asn Leu Lys Glu Ile Ile Val Ile Ala Asn Asp Asp Asp Ile Asn Ser  
 1460 1465 1470  
 Leu Ile Thr Glu Phe Leu Thr Leu Asp Ile Leu Val Phe Leu Lys Thr  
 1475 1480 1485  
 Phe Gly Gly Leu Leu Val Asn Gln Phe Ala Tyr Thr Leu Tyr Ser Leu  
 1490 1495 1500  
 Lys Ile Glu Gly Arg Asp Leu Ile Trp Asp Tyr Ile Met Arg Thr Leu  
 1505 1510 1515 1520  
 Arg Asp Thr Ser His Ser Ile Leu Lys Val Leu Ser Asn Ala Leu Ser  
 1525 1530 1535  
 His Pro Lys Val Phe Lys Arg Phe Trp Asp Cys Gly Val Leu Asn Pro  
 1540 1545 1550  
 Ile Tyr Gly Pro Asn Ile Ala Ser Gln Asp Gln Ile Lys Leu Ala Leu  
 1555 1560 1565

SUBSTITUTE SHEET (RULE 26)

- 271 -

Ser Ile Cys Glu Tyr Ser Leu Asp Leu Phe Met Arg Glu Trp Leu Asn  
 1570 1575 1580

Gly Val Ser Leu Glu Ile Tyr Ile Cys Asp Ser Asp Met Glu Val Ala  
 1585 1590 1595 1600

Asn Asp Arg Lys Gln Ala Phe Ile Ser Arg His Leu Ser Phe Val Cys  
 1605 1610 1615

Cys Leu Ala Glu Ile Ala Ser Phe Gly Pro Asn Leu Leu Asn Leu Thr  
 1620 1625 1630

Tyr Leu Glu Arg Leu Asp Leu Leu Lys Gln Tyr Leu Glu Leu Asn Ile  
 1635 1640 1645

Lys Glu Asp Pro Thr Leu Lys Tyr Val Gln Ile Ser Gly Leu Leu Ile  
 1650 1655 1660

Lys Ser Phe Pro Ser Thr Val Thr Tyr Val Arg Lys Thr Ala Ile Lys  
 1665 1670 1675 1680

Tyr Leu Arg Ile Arg Gly Ile Ser Pro Pro Glu Val Ile Asp Asp Trp  
 1685 1690 1695

Asp Pro Val Glu Asp Glu Asn Met Leu Asp Asn Ile Val Lys Thr Ile  
 1700 1705 1710

Asn Asp Asn Cys Asn Lys Asp Asn Lys Gly Asn Lys Ile Asn Asn Phe  
 1715 1720 1725

Trp Gly Leu Ala Leu Lys Asn Tyr Gln Val Leu Lys Ile Arg Ser Ile  
 1730 1735 1740

Thr Ser Asp Ser Asp Asp Asn Asp Arg Leu Asp Ala Asn Thr Ser Gly  
 1745 1750 1755 1760

Leu Thr Leu Pro Gln Gly Gly Asn Tyr Leu Ser His Gln Leu Arg Leu  
 1765 1770 1775

Phe Gly Ile Asn Ser Thr Ser Cys Leu Lys Ala Leu Glu Leu Ser Gln  
 1780 1785 1790

Ile Leu Met Lys Glu Val Asn Lys Asp Lys Asp Arg Leu Phe Leu Gly  
 1795 1800 1805

Glu Gly Ala Gly Ala Met Leu Ala Cys Tyr Asp Ala Thr Leu Gly Pro  
 1810 1815 1820

Ala Val Asn Tyr Tyr Asn Ser Gly Leu Asn Ile Thr Asp Val Ile Gly  
 1825 1830 1835 1840

Gln Arg Glu Leu Lys Ile Phe Pro Ser Glu Val Ser Leu Val Gly Lys

SUBSTITUTE SHEET (RULE 26)

- 272 -

1845	1850	1855
Lys Leu Gly Asn Val Thr Gln Ile Leu Asn Arg Val Lys Val Leu Phe 1860	1865	1870
Asn Gly Asn Pro Asn Ser Thr Trp Ile Gly Asn Met Glu Cys Glu Ser 1875	1880	1885
Leu Ile Trp Ser Glu Leu Asn Asp Lys Ser Ile Gly Leu Val His Cys 1890	1895	1900
Asp Met Glu Gly Ala Ile Gly Lys Ser Glu Glu Thr Val Leu His Glu 1905	1910	1915 1920
His Tyr Ser Val Ile Arg Ile Thr Tyr Leu Ile Gly Asp Asp Asp Val 1925	1930	1935
Val Leu Val Ser Lys Ile Ile Pro Thr Ile Thr Pro Asn Trp Ser Arg 1940	1945	1950
Ile Leu Tyr Leu Tyr Lys Leu Tyr Trp Lys Asp Val Ser Ile Ile Ser 1955	1960	1965
Leu Lys Thr Ser Asn Pro Ala Ser Thr Glu Leu Tyr Leu Ile Ser Lys 1970	1975	1980
Asp Ala Tyr Cys Thr Ile Met Glu Pro Ser Glu Ile Val Leu Ser Lys 1985	1990	1995 2000
Leu Lys Arg Leu Ser Leu Leu Glu Glu Asn Asn Leu Leu Lys Trp Ile 2005	2010	2015
Ile Leu Ser Lys Lys Arg Asn Asn Glu Trp Leu His His Glu Ile Lys 2020	2025	2030
Glu Gly Glu Arg Asp Tyr Gly Ile Met Arg Pro Tyr His Met Ala Leu 2035	2040	2045
Gln Ile Phe Gly Phe Gln Ile Asn Leu Asn His Leu Ala Lys Glu Phe 2050	2055	2060
Leu Ser Thr Pro Asp Leu Thr Asn Ile Asn Asn Ile Ile Gln Ser Phe 2065	2070	2075 2080
Gln Arg Thr Ile Lys Asp Val Leu Phe Glu Trp Ile Asn Ile Thr His 2085	2090	2095
Asp Asp Lys Arg His Lys Leu Gly Gly Arg Tyr Asn Ile Phe Pro Leu 2100	2105	2110
Lys Asn Lys Gly Lys Leu Arg Leu Leu Ser Arg Arg Leu Val Leu Ser 2115	2120	2125

SUBSTITUTE SHEET (RULE 26)



- 273 -

Trp Ile Ser Leu Ser Leu Ser Thr Arg Leu Leu Thr Gly Arg Phe Pro  
2130 2135 2140

Asp Glu Lys Phe Glu His Arg Ala Gln Thr Gly Tyr Val Ser Leu Ala  
2145 2150 2155 2160

Asp Thr Asp Leu Glu Ser Leu Lys Leu Leu Ser Lys Asn Ile Ile Lys  
2165 2170 2175

Asn Tyr Arg Glu Cys Ile Gly Ser Ile Ser Tyr Trp Phe Leu Thr Lys  
2180 2185 2190

Glu Val Lys Ile Leu Met Lys Leu Ile Gly Gly Ala Lys Leu Leu Gly  
2195 2200 2205

Ile Pro Arg Gln Tyr Lys Glu Pro Glu Asp Gln Leu Leu Glu Asn Tyr  
2210 2215 2220

Asn Gln His Asp Glu Phe Asp Ile Asp  
2225 2230

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15462 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACCAAACAAG AGAAGAACT TGCTTGGTAA TATAAATTTA ACTTAAAATT AACTTAGGAT 60  
TTAAGACATT GACTAGAAGG TCAAGAAAAG GGAAGCTCTAT AATTTCAAAA ATGTTGAGCC 120  
TATTTGATAC ATTTAATGCA CGTAGGCAAG AAAACATAAC AAAATCAGCC GGTGGAGCTA 180  
TCATTCCCTGG ACAGAAAAAT ACTGTCTCTA TATTCGCCCT TGGACCGACA ATAAGTGATG 240  
ATAATGAGAA AATGACATTA GCTCTTCTAT TTCTATCTCA TTTACTAGAT AATGAGAAAC 300  
AACATGCACA AAGGGCAGGG TTCTTGGTGT CTTTATTGTC AATGGCTTAT GCCAATCCAG 360  
AGCTCTACCT AACAACAAAT GGAAGTAATG CAGATGCCAA GTATGTCATA TACATGATTG 420  
AGAAAGATCT AAAACGGCAA AAGTATGGAG GATTTGTGGT TAAGACGAGA GAGATGATAT 480

SUBSTITUTE SHEET (RULE 26)

- 274 -

ATGAAAAGAC AACTGATTGG ATATTTGGAA GTGACCTGGA TTATGATCAG GAACTATGT	540
TGCAGAACGG CAGGAACAAT TCAACAATTG AAGACCTTGT CCACACATTG GGGTATCCAT	600
CATGTTTAGG AGCTCTTATA ATACAGATCT GGATAGTTCT GGTCAAAGCT ATCACTAGTA	660
TCTCAGGGTT AAGAAAAGGC TTTTTCACCC GATTGGAAGC TTTCAGACAA GATGGAACAG	720
TGCAGGCAGG GCTGGTATTG AGCGGTGACA CAGTGGATCA GATTGGGTCA ATCATGCGGT	780
CTCAACAGAG CTTGGTAACT CTTATGGTTG AAACATTAAT AACATGAAT ACCAGCAGAA	840
ATGACCTCAC AACCATAGAA AAGAATATAC AAATTGTTGG CAACTACATA AGAGATGCAG	900
GTCTCGCTTC ATTCTTCAAT ACAATCAGAT ATGGAATTGA GACCAGAATG GCAGCTTTGA	960
CTCTATCCAC TCTCAGACCA GATATCAATA GATTAAAAGC TTTGATGGAA CTGTATTTAT	1020
CAAAGGGACC ACGCGCTCCT TTCATCTGTA TCCTCAGAGA TCCTATACAT GGTGAGTTCTG	1080
CACCAGGCAA CTATCCTGCC ATATGGAGCT ATGCAATGGG GGTGGCAGTT GTACAAAATA	1140
GAGCCATGCA ACAGTATGTG ACGGGAAGAT CATATCTAGA CATTGATATG TTCCAGCTAG	1200
GACAAGCAGT AGCACGTGAT GCCGAAGCTC AAATGAGCTC AACACTGGAA GATGAACTTG	1260
GAGTGACACA CGAAGCTAAA GAAAGCTTGA AGAGACATAT AAGGAACATA AACAGTTCAG	1320
AGACATCTTT CCACAAACCG ACAGGTGGAT CAGCCATAGA GATGGCAATA GATGAAGAGC	1380
CAGAACAATT CGAACATAGA GCAGATCAAG AACAAAATGG AGAACCTCAA TCATCCATAA	1440
TTCAATATGC CTGGGCAGAA GGAAATAGAA GCGATGATCA GACTGAGCAA GCTACAGAAT	1500
CTGACAATAT CAAGACCGAA CAACAAAACA TCAGAGACAG ACTAAACAAG AGACTCAACG	1560
ACAAGAAGAA ACAAGCAGT CAACCACCCA CTAATCCCAC AACAGAACA AACCAGGACG	1620
AAATAGATGA TCTGTTTAAC GCATTTGGAA GCAACTAATC GAATCAACAT TTTAATCTAA	1680
ATCAATAATA AATAAGAAAA ACTTAGGATT AAAGAATCCT ATCATACCGG AATATAGGGT	1740
GGTAAATTTA GAGTCTGCTT GAAACTCAAT CAATAGAGAG TTGATGGAAA GCGATGCTAA	1800
AAACTATCAA ATCATGGATT CTTGGGAAGA GGAATCAAGA GATAAATCAA CTAATATCTC	1860
CTCGGCCCTC AACATCATTG AATTCATACT CAGCACCGAC CCCCAGAAG ACTTATCGGA	1920
AAACGACACA ATCAACACAA GAACCCAGCA ACTCAGTGCC ACCATCTGTC AACCAGAAAT	1980
CAAACCAACA GAAACAAGTG AGAAAGATAG TGGATCAACT GACAAAAATA GACAGTCTGG	2040

SUBSTITUTE SHEET (RULE 26)

- 275 -

GTCATCACAC GAATGTACAA CAGAAGCAAA AGATAGAAAC ATTGATCAGG AACTGTACA 2100  
GAGAGGACCT GGGAGAAGAA GCAGCTCAGA TAGTAGAGCT GAGACTGTGG TCTCTGGAGG 2160  
AATCCCCAGA AGCATCACAG ATTCTAAAAA TGGAAACCAA AACACGGAGG ATATTGATCT 2220  
CAATGAAATT AGAAAGATGG ATAAGGACTC TATTGAGGGG AAAATGCGAC AATCTGCAAA 2280  
TGTTCCAAGC GAGATATCAG GAAGTGATGA CATATTTACA ACAGAACAAA GTAGAAACAG 2340  
TGATCATGGA AGAAGCCTGG AATCTATCAG TACACCTGAT ACAAGATCAA TAAGTGTGT 2400  
TACTGCTGCA ACACCAGATG ATGAAGAAGA AATACTAATG AAAAATAGTA GGACAAAGAA 2460  
AAGTTCTTCA ACACATCAAG AAGATGACAA AAGAATTAAA AAAGGGGGAA AAGGGAAAGA 2520  
CTGGTTTAAG AAATCAAAAG ATACCGACAA CCAGATACCA ACATCAGACT ACAGATCCAC 2580  
ATCAAAAGGG CAGAAGAAAA TCTCAAAGAC AACACCACC AACACCGACA CAAAGGGGCA 2640  
AACAGAAATA CAGACAGAAT CATCAGAAAC ACAATCCTCA TCATGGAATC TCATCATCGA 2700  
CAACAACACC GACCGGAACG AACAGACAAG CACAACCTCT CCAACAACAA CTTCCAGATC 2760  
AACTTATACA AAAGAATCGA TCCGAACAAA CTCTGAATCC AAACCCAAGA CACAAAAGAC 2820  
AAATGGAAAG GAAAGGAAGG ATACAGAAGA GAGCAATCGA TTTACAGAGA GGGCAATTAC 2880  
TCTATTGCAG AATCTTGGTG TAATTCAATC CACATCAAAA CTAGATTAT ATCAAGACAA 2940  
ACGAGTTGTA TGTGTAGCAA ATGTACTAAA CAATGTAGAT ACTGCATCAA AGATAGATT 3000  
CCTGGCAGGA TTAGTCATAG GGGTTTCAAT GGACAACGAC ACAAATTAAC CACAGATACA 3060  
AAATGAAATG CTAAACCTCA AAGCAGATCT AAAGAAAATG GACGAATCAC ATAGAAGATT 3120  
GATAGAAAAT CAAAGAGAAC AACTGTCATT GATCACGTCA CTAATTTCAA ATCTCAAAT 3180  
TATGACTGAG AGAGGAGGAA AGAAAGACCA AAATGAATCC AATGAGAGAG TATCCATGAT 3240  
CAAAACAAAA TTGAAAGAAG AAAAGATCAA GAAGACCAGG TTTGACCCAC TTATGGAGGC 3300  
ACAAGGCATT GACAAGAATA TACCCGATCT ATATCGACAT GCAGGAGATA CACTAGAGAA 3360  
CGATGTACAA GTTAAATCAG AGATATTAAG TTCATACAAT GAGTCAAATG CAACAAGACT 3420  
AATACCCAAA AAAGTGAGCA GTACAATGAG ATCACTAGTT GCAGTCATCA ACAACAGCAA 3480  
TCTCTCACAA AGCACAAAAC AATCATACAT AAACGAACTC AAACGTTGCA AAAATGATGA 3540  
AGAAGTATCT GAATTAATGG ACATGTTCAA TGAAGATGTC AACAAATGCC AATGATCCAA 3600

SUBSTITUTE SHEET (RULE 26)

- 276 -

CAAAGAAACG ACACCGAACA AACAGACAAG AAACAACAGT AGATCAAAAC CTGTCAACAC	3660
ACACAAAATC AAGCAGAATG AAACAACAGA TATCAATCAA TATACAAATA AGAAAACTT	3720
AGGATTAAAG AATAAATTAA TCCTTGTTCA AAATGAGTAT AACTAACTCT GCAATATACA	3780
CATTCCCAGA ATCATCATTC TCTGAAAATG GTCATATAGA ACCATTACCA CTCAAAGTCA	3840
ATGAACAGAG GAAAGCAGTA CCCCACATTA GAGTTGCCAA GATCGGAAAT CCACCAAAAC	3900
ACGGATCCCG GTATTTAGAT GTCTTCTTAC TCGGCTTCTT CGAGATGGAA CGAATCAAAG	3960
ACAAATACGG GAGTGTGAAT GATCTCGACA GTGACCCGAG TTACAAAGTT TGTGGCTCTG	4020
GATCATTACC AATCGGATTG GCTAAGTACA CTGGGAATGA CCAGGAATTG TTACAAGCCG	4080
CAACCAAACCT GGATATAGAA GTGAGAAGAA CAGTCAAAGC GAAAGAGATG GTTGTTTACA	4140
CGGTACAAAA TATAAAACCA GAACTGTACC CATGGTCCAA TAGACTAAGA AAAGGAATGC	4200
TGTTGATGTC CAACAAAGTT GCTCTTGCTC CTCAATGTCT TCCACTAGAT AGGAGCATAA	4260
AATTTAGAGT AATCTTCGTG AATTGTACGG CAATTGGATC AATAACCTTG TTCAAATTC	4320
CTAAGTCAAT GGCATCACTA TCTCTAACCA ACACAATATC AATCAATCTG CAGGTACACA	4380
TAAAAACAGG GGTTCAGACT GATTCTAAAG GGATAGTTCA AATTTTGGAT GAGAAAGGCG	4440
AAAAATCACT GAATTTTCATG GTCCATCTCG GATTGATCAA AAGAAAAGTA GGCAGAATGT	4500
ACTCTGTTGA ATACTGTAAA CAGAAAATCG AGAAAATGAG ATTGATATTT TCTTTAGGAC	4560
TAGTTGGAGG AATCAGTCTT CATGTCAATG CAACTGGGTC CATATCAAAA AACTAGCAA	4620
GTCAGCTGGT ATTCAAAAGA GAGATTTGTT ATCCTTTAAT GGATCTAAAT CCGCATCTCA	4680
ATCTAGTTAT CTGGGCTTCA TCAGTAGAGA TTACAAGAGT GGATGCAATT TTCCAACCTT	4740
CTTTACCTGG CGAGTTCAGA TACTATCCTA ATATTATTGC AAAAGGAGTT GGGAAAATCA	4800
AACAATGGAA CTAGTAATCT CTATTTTAGT CCGGACGTAT CTATTAAGCC GAAGCAAATA	4860
AAGGATAATC AAAAAGTTAG GACAAAAGAG GTCAATACCA ACAACTATTA GCAGTCACAC	4920
TCGCAAGAAT AAGAGAGAAG GGACCAAAAA AGTCAAATAG GAGAAATCAA AACAAAAGGT	4980
ACAGAACACC AGAACAACAA AATCAAAACA TCCAACCTAC TCAAAACAAA AATTCCAAAA	5040
GAGACCGGCA ACACAACAAG CACTGAACAC AATGCCAACT TCAATACTGC TAATTATTAC	5100
AACCATGATC ATGGCATCTT TCTGCCAAAT AGATATCACA AACTACAGC ACGTAGGTGT	5160

SUBSTITUTE SHEET (RULE 26)

- 277 -

ATTGGTCAAC AGTCCCAAAG GGATGAAGAT ATCACAAAAC TTTGAAACAA GATATCTAAT 5220  
TTTGAGCCTC ATACCAAAAA TAGAAGACTC TAACTCTTGT GGTGACCAAC AGATCAAGCA 5280  
ATACAAGAAG TTATTGGATA GACTGATCAT CCCTTTATAT GATGGATTAA GATTACAGAA 5340  
AGATGTGATA GTAACCAATC AAGAATCCAA TGAAACACT GATCCCAGAA CAAAACGATT 5400  
CTTTGGAGGG GTAATTGGAA CCATTGCTCT GGGAGTAGCA ACCTCAGCAC AAATTACAGC 5460  
GGCAGTTGCT CTGGTTGAAG CCAAGCAGGC AAGATCAGAC ATCGAAAAAC TCAAAGAAGC 5520  
AATTAGGGAC ACAAATAAAG CAGTGCAGTC AGTTCAGAGC TCCATAGGAA ATTTAATAGT 5580  
AGCAATTAAA TCAGTCCAGG ATTATGTTAA CAAAGAAATC GTGCCATCGA TTGCGAGGCT 5640  
AGGTTGTGAA GCAGCAGGAC TTCAATTAGG AATTGCATTA ACACAGCATT ACTCAGAATT 5700  
AACAAACATA TTTGGTGATA ACATAGGATC GTTACAAGAA AAAGGAATAA AATTACAAGG 5760  
TATAGCATCA TTATACCGCA CAAATATCAC AGAAATATTC ACAACATCAA CAGTTGATAA 5820  
ATATGATATC TATGATCTGT TATTACAGA ATCAATAAAG GTGAGAGTTA TAGATGTTGA 5880  
CTTGAATGAT TACTCAATCA CCCTCCAAGT CAGACTCCCT TTATTAACTA GGCTGCTGAA 5940  
CACTCAGATC TACAAAGTAG ATTCCATATC ATATAACATC CAAAACAGAG AATGGTATAT 6000  
CCCTCTTCCC AGCCATATCA TGACGAAAGG GGCATTTCTA GGTGGAGCAG ACGTCAAAGA 6060  
ATGTATAGAA GCATTCAGCA GCTATATATG CCCTTCTGAT CCAGGATTG TATTAAACCA 6120  
TGAAATAGAG AGCTGCTTAT CAGGAAACAT ATCCCAATGT CCAAGAACAA CGGTCACATC 6180  
AGACATTGTT CCAAGATATG CATTTGTCAA TGGAGGAGTG GTTGCAAACGT GTATAACAAC 6240  
CACCTGTACA TGCAACGGAA TTGGTAATAG AATCAATCAA CCACCTGATC AAGGAGTAAA 6300  
AATTATAACA CATAAAGAAT GTAGTACAGT AGGTATCAAC GGAATGCTGT TCAATACAAA 6360  
TAAAGAAGGA ACTCTTGCAT TCTATACACC AAATGATATA AACTAACA ATTCTGTTAC 6420  
ACTTGATCCA ATTGACATAT CAATCGAGCT CAACAAGGCC AAATCAGATC TAGAAGAATC 6480  
AAAAGAATGG ATAAGAAGGT CAAATCAAAA ACTAGATTCT ATTGGAAATT GGCATCAATC 6540  
TAGCACTACA ATCATAATTA TTTTGATAAT GATCATTATA TTGTTTATAA TTAATATAAC 6600  
GATAATTACA ATTGCAATTA AGTATTACAG AATTCAAAAG AGAAATCGAG TGGATCAAAA 6660  
TGACAAGCCA TATGTACTAA CAAACAAATA ACATATCTAC AGATCATTAG ATATTAAAAT 6720

SUBSTITUTE SHEET (RULE 26)

- 278 -

TATAAAAAAC	TTAGGAGTAA	AGTTACGCAA	TCCAACCTCTA	CTCATATAAT	TGAGGAAGGA	6780
CCCAATAGAC	AAATCCAAAT	TCGAGATGGA	ATACTGGAAG	CATACCAATC	ACGGAAAGGA	6840
TGCTGGCAAT	GAGCTGGAGA	CGTCTATGGC	TACTCATGGC	AACAAGCTCA	CTAATAAGAT	6900
AATATACATA	TTATGGACAA	TAATCCTGGT	GTTATTATCA	ATAGTCTTCA	TCATAGTGCT	6960
AATTAATTCC	ATCAAAAGTG	AAAAGGCCCA	CGAATCATTG	CTGCAAGACA	TAAATAATGA	7020
GTTCATGGAA	ATTACAGAAA	AGATCCAAAT	GGCATCGGAT	AATACCAATG	ATCTAATACA	7080
GTCAGGAGTG	AATACAAGGC	TTCTTACAAT	TCAGAGTCAT	GTCCAGAATT	ACATACCAAT	7140
ATCATTGACA	CAACAGATGT	CAGATCTTAG	GAAATTCATT	AGTGAAATTA	CAATTAGAAA	7200
TGATAATCAA	GAAGTGCTGC	CACAAAGAAT	AACACATGAT	GTAGGTATAA	AACCTTTAAA	7260
TCCAGATGAT	TTTTGGAGAT	GCACGTCTGG	TCTTCCATCT	TTAATGAAAA	CTCCAAAAAT	7320
AAGGTTAATG	CCAGGGCCGG	GATTATTAGC	TATGCCAACG	ACTGTTGATG	GCTGTGTTAG	7380
AACTCCGTCT	TTAGTTATAA	ATGATCTGAT	TTATGCTTAT	ACCTCAAATC	TAATTACTCG	7440
AGGTTGTCAG	GATATAGGAA	AATCATATCA	AGTCTTACAG	ATAGGGATAA	TAACTGTAAA	7500
CTCAGACTTG	GTACCTGACT	TAAATCCTAG	GATCTCTCAT	ACCTTTAACA	TAAATGACAA	7560
TAGGAAGTCA	TGTTCTCTAG	CACTCCTAAA	TACAGATGTA	TATCAACTGT	GTTCAACTCC	7620
CAAAGTTGAT	GAAAGATCAG	ATTATGCATC	ATCAGGCATA	GAAGATATTG	TACTTGATAT	7680
TGTCAATTAT	GATGGTTCAA	TCTCAACAAC	AAGATTTAAG	AATAATAACA	TAAGCTTTGA	7740
TCAACCATAT	GCTGCACTAT	ACCCATCTGT	TGGACCAGGG	ATATACTACA	AAGGCAAAAT	7800
AATATTTCTC	GGGTATGGAG	GTCTTGAACA	TCCAATAAAT	GAGAATGTAA	TCTGCAACAC	7860
AACTGGGTGC	CCCGGGAAAA	CACAGAGAGA	CTGTAATCAA	GCGTCTCATA	GTCCATGGTT	7920
TTCAGATAGG	AGGATGGTCA	ACTCCATCAT	TGTTGCTGAC	AAAGGCTTAA	ACTCAATTCC	7980
AAAATTGAAA	GTATGGACGA	TATCTATGCG	ACAAAATTAC	TGGGGGTCAG	AAGGAAGGTT	8040
ACTTCTACTA	GGTAACAAGA	TCTATATATA	TACAAGATCT	ACAAGTTGGC	ATAGCAAGTT	8100
ACAATTAGGA	ATAATTGATA	TTACTGATTA	CAGTGATATA	AGGATAAAAT	GGACATGGCA	8160
TAATGTGCTA	TCAAGACCAG	GAAACAATGA	ATGTCCATGG	GGACATTCAT	GTCCAGATGG	8220
ATGTATAACA	GGAGTATATA	CTGATGCATA	TCCAATCAAT	CCCACAGGGA	GCATTGTGTC	8280

SUBSTITUTE SHEET (RULE 26)

- 279 -

ATCTGTCATA TTAGACTCAC AAAAATCGAG AGTGAACCCA GTCATAACTT ACTCAACAGC	8340
AACCGAAAGA GTAAACGAGC TGGCCATCCT AAACAGAACA CTCTCAGCTG GATATACAAC	8400
AACAAGCTGC ATTACACACT ATAACAAAGG ATATTGTTTT CATATAGTAG AAATAAATCA	8460
TAAAAGCTTA AACACATTTT AACCCATGTT GTTCAAAACA GAGATTCCAA AAAGCTGCAG	8520
TTAATCATAA TTAACCATAA TATGCATCAA TCTATCTATA ATACAAGTAT ATGATAAGTA	8580
ATCAGCAATC AGACAATAGA CAAAAGGGAA ATATAAAAAA CTTAGGAGCA AAGCGTGCTC	8640
GGGAAATGGA CACTGAATCT AACAATGGCA CTGTATCTGA CATACTCTAT CCTGAGTGTC	8700
ACCTTAACTC TCCTATCGTT AAAGGTAAAA TAGCACAATT ACACACTATT ATGAGTCTAC	8760
CTCAGCCTTA TGATATGGAT GACGACTCAA TACTAGTTAT CACTAGACAG AAAATAAAAC	8820
TTAATAAATT GGATAAAAGA CAACGATCTA TTAGAAGATT AAAATTAATA TTAAGTAAA	8880
AAGTGAATGA CTTAGGAAAA TACACATTTA TCAGATATCC AGAAATGTCA AAAGAAATGT	8940
TCAAATTATA TATACCTGGT ATTAACAGTA AAGTGACTGA ATTATTACTT AAAGCAGATA	9000
GAACATATAG TCAAATGACT GATGGATTAA GAGATCTATG GATTAATGTG CTATCAAAAT	9060
TAGCCTCAAA AAATGATGGA AGCAATTATG ATCTTAATGA AGAAATTAAT AATATATCGA	9120
AAGTTCACAC AACCTATAAA TCAGATAAAT GGTATAATCC ATTCAAAACA TGGTTTACTA	9180
TCAAGTATGA TATGAGAAGA TTACAAAAAG CTCGAAATGA GATCACTTTT AATGTTGGGA	9240
AGGATTATAA CTTGTTAGAA GACCAGAAGA ATTTCTTATT GATACATCCA GAATTGGTTT	9300
TGATATTAGA TAAACAAAAC TACAATGGTT ATCTAATTAC TCCTGAATTA GTATTGATGT	9360
ATTGTGACGT AGTCGAAGGC CGATGGAATA TAAGTGCATG TGCTAAGTTA GATCCAAAAT	9420
TACAATCTAT GTATCAGAAA GGTAATAACC TGTGGGAAGT GATAGATAAA TTGTTTCCAA	9480
TTATGGGAGA AAAGACATTT GATGTGATAT CGTTATTAGA ACCACTTGCA TTATCCTTAA	9540
TTCAAACCTCA TGATCCTGTT AAACAACATA GAGGAGCTTT TTAAATCAT GTGTTATCCG	9600
AGATGGAATT AATATTGAA TCTAGAGAAT CGATTAAGGA ATTTCTGAGT GTAGATTACA	9660
TTGATAAAAT TTTAGATATA TTTAATAAGT CTACAATAGA TGAAATAGCA GAGATTTTCT	9720
CTTTTTTTAG AACATTGGG CATCCTCCAT TAGAAGCTAG TATTGCAGCA GAAAAGGTTA	9780
GAAAATATAT GTATATTGGA AAACAATTAA AATTGACAC TATTAATAAA TGTCATGCTA	9840

SUBSTITUTE SHEET (RULE 26)

- 280 -

TCTTCTGTAC AATAATAATT AACGGATATA GAGAGAGGCA TGGTGGACAG TGGCCTCCTG 9900  
TGACATTACC TGATCATGCA CACGAATTCA TCATAAATGC TTACGGTTCA AACTCTGCGA 9960  
TATCATATGA GAATGCTGTT GATTATTACC AGAGCTTTAT AGGAATAAAA TTCAATAAAT 10020  
TCATAGAGCC TCAGTTAGAT GAGGATTTGA CAATTTATAT GAAAGATAAA GCATTATCTC 10080  
CAAAAAAATC AAATTGGGAC ACAGTTTATC CTGCATCTAA TTTACTGTAC CGTACTAACG 10140  
CATCCAACGA ATCACGAAGA TTAGTTGAAG TATTTATAGC AGATAGTAAA TTTGATCCTC 10200  
ATCAGATATT GGATTATGTA GAATCTGGGG ACTGGTTAGA TGATCCAGAA TTTAATATTT 10260  
CTTATAGTCT TAAAGAAAAA GAGATCAAAC AGGAAGGTAG ACTCTTTGCA AAAATGACAT 10320  
ACAAAATGAG AGCTACACAA GTTTTATCAG AGACACTACT TGCAAATAAC ATAGGAAAAAT 10380  
TCTTCAAGA AAATGGGATG GTGAAGGGAG AGATTGAATT ACTTAAGAGA TTAACAACCA 10440  
TATCAATATC AGGAGTTCCA CGGTATAATG AAGTGTACAA TAATTCTAAA AGCCATACAG 10500  
ATGACCTTAA AACCTACAAT AAAATAAGTA ATCTTAATTT GTCTTCTAAT CAGAAATCAA 10560  
AGAAATTTGA ATTCAAGTCA ACGGATATCT ACAATGATGG ATACGAGACT GTGAGCTGTT 10620  
TCCTAACAAC AGATCTCAA AAATACTGTC TTAATTGGAG ATATGAATCA ACAGCTCTAT 10680  
TTGAGAAAC TTGCAACCAA ATATTTGGAT TAAATAAATT GTTTAATTGG TTACACCCTC 10740  
GTCTGAAGG AAGTACAATC TATGTAGGTG ATCCTTACTG TCCTCCATCA GATAAAGAAC 10800  
ATATATCATT AGAGGATCAC CCTGATTCTG GTTTTTACGT TCATAACCCA AGAGGGGGTA 10860  
TAGAAGGATT TTGTCAAAAA TTATGGACAC TCATATCTAT AAGTGCAATA CATCTAGCAG 10920  
CTGTTAGAAT AGGCGTGAGG GTGACTGCAA TGGTTCAAGG AGACAATCAA GCTATAGCTG 10980  
TAACCACAAG AGTACCCAAC AATTATGACT ACAGAGTTAA GAAGGAGATA GTTTATAAAG 11040  
ATGTAGTGAG ATTTTTTGAT TCATTAAGAG AAGTGATGGA TGATCTAGGT CATGAACTTA 11100  
AATTAAATGA AACGATTATA AGTAGCAAGA TGTTCAATA TAGCAAAGA ATCTATTATG 11160  
ATGGGAGAAT TCTTCTCAA GCTCTAAAAG CATTATCTAG ATGTGTCTTC TGGTCAGAGA 11220  
CAGTAATAGA CGAAACAAGA TCAGCATCTT CAAATTTGGC AACATCATTT GCAAAAGCAA 11280  
TTGAGAATGG TTATTCACCT GTTCTAGGAT ATGCATGCTC AATTTTAAAG AACATTCAAC 11340  
AACTATATAT TGCCCTTGGG ATGAATATCA ATCCAATAT AACACAGAAT ATCAGAGATC 11400

SUBSTITUTE SHEET (RULE 26)



- 281 -

AGTATTTTAG	GAATCCAAAT	TGGATGCAAT	ATGCCTCTTT	AATACCTGCT	AGTGTGTTGGG	11460
GATTCAATCA	CATGGCCATG	TCAAGATGTT	TTGTAAGGAA	TATTGGTGAT	CCATCAGTTG	11520
CCGCATTGGC	TGATATTAAA	AGATTTATTA	AGGCGAATCT	ATTAGACCGA	AGTGTCTCTT	11580
ATAGGATTAT	GAATCAAGAA	CCAGGTGAGT	CATCTTTTTT	TGACTGGGCT	TCAGATCCAT	11640
ATTCATGCAA	TTTACCACAA	TCTCAAATA	TAACCACCAT	GATAAAAAAT	ATAACAGCAA	11700
GGAATGTATT	ACAAGATTCA	CCAAATCCAT	TATTATCTGG	ATTATTCACA	AATACAATGA	11760
TAGAAGAAGA	TGAAGAATTA	GCTGAGTTCC	TGATGGACAG	GAAGGTAATT	CTCCCTAGAG	11820
TTGCACATGA	TATTCTAGAT	AATTCTCTCA	CAGGAATTAG	AAATGCCATA	GCTGGAATGT	11880
TAGATACGAC	AAAATCACTA	ATTCGGGTTG	GCATAAATAG	AGGAGGACTG	ACATATAGTT	11940
TGTTGAGGAA	AATCAGTAAT	TACGATCTAG	TACAATATGA	AACACTAAGT	AGGACTTTGC	12000
GACTAATTGT	AAGTGATAAA	ATCAAGTATG	AAGATATGTG	TTCGGTAGAC	CTTGCCATAG	12060
CATTGCGACA	AAAGATGTGG	ATTCATTTAT	CAGGAGGAAG	GATGATAAGT	GGACTTGAAA	12120
CGCCTGACCC	ATTAGAATTA	CTATCTGGGG	TAGTAATAAC	AGGATCAGAA	CATTGTAAAA	12180
TATGTTATTC	TTCAGATGGC	ACAAACCCAT	ATACTTGGAT	GTATTTACCC	GGTAATATCA	12240
AAATAGGATC	AGCAGAAACA	GGTATATCGT	CATTAAGAGT	TCCTTATTTT	GGATCAGTCA	12300
CTGATGAAAG	ATCTGAAGCA	CAATTAGGAT	ATATCAAGAA	TCTTAGTAAA	CCTGCAAAAG	12360
CCGCAATAAG	AATAGCAATG	ATATATACAT	GGGCATTGG	TAATGATGAG	ATATCTTGGA	12420
TGGAAGCCTC	ACAGATAGCA	CAAACACGTG	CAAATTTTAC	ACTAGATAGT	CTCAAAATTT	12480
TAACACCGGT	AGCTACATCA	ACAAATTTAT	CACACAGATT	TAAGGATACT	GCAACTCAGA	12540
TGAAATTCTC	CAGTACATCA	TTGATCAGAG	TCAGCAGATT	TATAACAATG	TCCAATGATA	12600
ACATGTCTAT	CAAAGAAGCT	AATGAAACCA	AAGATACTAA	TCTTATTTAT	CAACAAATAA	12660
TGTTAACAGG	ATTAAGTGTT	TTCGAATATT	TATTTAGATT	AAAAGAAACC	ACAGGACACA	12720
ACCCTATAGT	TATGCATCTG	CACATAGAAG	ATGAGTGTG	TATTAAAGAA	AGTTTAAATG	12780
ATGAACATAT	TAATCCAGAG	TCTACATTAG	AATTAATTG	ATATCCTGAA	AGTAATGAAT	12840
TTATTTATGA	TAAAGACCCA	CTCAAAGATG	TGGACTTATC	AAAACCTATG	GTTATTAAAG	12900
ACCATTCTTA	CACAATTGAT	ATGAATTATT	GGGATGATAC	TGACATCATA	CATGCAATTT	12960

SUBSTITUTE SHEET (RULE 26)

- 282 -

CAATATGTAC TGCAATTACA ATAGCAGATA CTATGTCACA ATTAGATCGA GATAATTTAA	13020
AAGAGATAAT AGTTATTGCA AATGATGATG ATATTAATAG CTTAATCACT GAATTTTGA	13080
CTCTTGACAT ACTTGATTTT CTCAAGACAT TTGGTGGATT ATTAGTAAAT CAATTTGCAT	13140
ACACTCTTTA TAGTCTAAAA ATAGAAGGTA GGGATCTCAT TTGGGATTAT ATAATGAGAA	13200
CACTGAGAGA TACTTCCCAT TCAATATTAA AAGTATTATC TAATGCATTA TCTCATCCTA	13260
AAGTATTCAA GAGGTTCTGG GATTGTGGAG TTTAAACCC TATTATGGT CCTAATATTG	13320
CTAGTCAAGA CCAGATAAAA CTGCCCCTAT CTATATGTGA ATATTCACCTA GATCTATTTA	13380
TGAGAGAATG GTTGAATGGT GTATCACTTG AAATATACAT TTGTGACAGC GATATGGAAG	13440
TTGCAAATGA TAGGAAACAA GCCTTTATTT CTAGACACCT TTCATTTGTT TGTTGTTTAG	13500
CAGAAATTGC ATCTTTCGGA CCTAACCTGT TAACTTAAC ATACTTGGAG AGACTTGATC	13560
TATTGAAACA ATATCTTGAA TTAAATATTA AAGAAGACCC TACTCTTAA TATGTACAAA	13620
TATCTGGATT ATTAATTAAA TCGTTCCCAT CAACTGTAAC ATACGTAAGA AAGACTGCAA	13680
TCAAATATCT AAGGATTTCGC GGTATTAGTC CACCTGAGGT AATTGATGAT TGGGATCCGG	13740
TAGAAGATGA AAATATGCTG GATAACATTG TCAAACTAT AAATGATAAC TGTAATAAAG	13800
ATAATAAAGG GAATAAAATT AACAATTCT GGGGACTAGC ACTTAAGAAC TATCAAGTCC	13860
TTAAATCAG ATCTATAACA AGTGATTCTG ATGATAATGA TAGACTAGAT GCTAATACAA	13920
GTGGTTTGAC ACTTCCTCAA GGAGGGAATT ATCTATCGCA TCAATTGAGA TTATTCGGAA	13980
TCAACAGCAC TAGTTGTCTG AAAGCTCTG AGTTATCACA AATTTTAATG AAGGAAGTCA	14040
ATAAAGACAA GGACAGGCTC TTCCTGGGAG AAGGAGCAGG AGCTATGCTA GCATGTTATG	14100
ATGCCACATT AGGACCTGCA GTTAATTATT ATAATTCAGG TTTGAATATA ACAGATGTAA	14160
TTGGTCAACG AGAATTGAAA ATATTCCTT CAGAGGTATC ATTAGTAGGT 'AAAAAATTAG	14220
GAAATGTGAC ACAGATTCTT AACAGGGTAA AAGTACTGTT CAATGGGAAT CCTAATTCAA	14280
CATGGATAGG AAATATGGAA TGTGAGAGCT TAATATGGAG TGAATTAAAT GATAAGTCCA	14340
TTGGATTAGT ACATTGTGAT ATGGAAGGAG CTATCGGTAA ATCAGAAGAA ACTGTTCTAC	14400
ATGAACATTA TAGTGTTATA AGAATTACAT ACTTGATTGG GGATGATGAT GTTGTTTTAG	14460
TTTCCAAAAT TATACCTACA ATCACTCCGA ATTGGTCTAG AATACTTTAT CTATATAAAT	14520

- 283 -

TATATTGGAA AGATGTAAGT ATAATATCAC TCAAACTTC TAATCCTGCA TCAACAGAAT 14580  
 TATATCTAAT TTCGAAAGAT GCATATTGTA CTATAATGGA ACCTAGTGAA ATTGTTTTAT 14640  
 CAAAACCTAA AAGATTGTCA CTCTTGGAAG AAAATAATCT ATTAAATGG ATCATTTTAT 14700  
 CAAAGAAGAG GAATAATGAA TGGTTACATC ATGAAATCAA AGAAGGAGAA AGAGATTATG 14760  
 GAATCATGAG ACCATATCAT ATGGCACTAC AAATCTTTGG ATTTCAAATC AATTTAAATC 14820  
 ATCTGGCGAA AGAATTTTTA TCAACCCCAG ATCTGACTAA TATCAACAAT ATAATCCAAA 14880  
 GTTTCAGCG AACATAAAG GATGTTTTAT TTGAATGGAT TAATATAACT CATGATGATA 14940  
 AGAGACATAA ATTAGGCGGA AGATATAACA TATCCCACT GAAAAATAAG GGAAAGTTAA 15000  
 GACTGCTATC GAGAAGACTA GTATTAGTT GGATTTCATT ATCATTATCG ACTCGATTAC 15060  
 TTACAGGTCG CTTTCCTGAT GAAAAATTG AACATAGAGC ACAGACTGGA TATGTATCAT 15120  
 TAGCTGATAC TGATTTAGAA TCATTAAAGT TATTGTCGAA AAACATCATT AAGAATTACA 15180  
 GAGAGTGTAT AGGATCAATA TCATATTGGT TTCTAACCA AGAAGTTAAA ATACTTATGA 15240  
 AATTGATTGG TGGTGCTAAA TTATTAGGAA TTCCAGACA ATATAAAGAA CCCGAAGACC 15300  
 AGTTATTAGA AACTACAAT CAACATGATG AATTGATAT CGATTAAAC ATAAATACAA 15360  
 TGAAGATATA TCCTAACCTT TATCTTTAAG CCTAGGAATA GACAAAAAGT AAGAAAAACA 15420  
 TGTAATATAT ATATACCAA CAGAGTTCTT CTCTGTTTG GT 15462

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2233 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Asp Thr Glu Ser Asn Asn Gly Thr Val Ser Asp Ile Leu Tyr Pro  
 1 5 10 15  
 Glu Cys His Leu Asn Ser Pro Ile Val Lys Gly Lys Ile Ala Gln Leu

SUBSTITUTE SHEET (RULE 26)

- 284 -

20	25	30
His Thr Ile Met Ser Leu Pro Gln Pro Tyr Asp Met Asp Asp Asp Ser		
35	40	45
Ile Leu Val Ile Thr Arg Gln Lys Ile Lys Leu Asn Lys Leu Asp Lys		
50	55	60
Arg Gln Arg Ser Ile Arg Arg Leu Lys Leu Ile Leu Thr Glu Lys Val		
65	70	75 80
Asn Asp Leu Gly Lys Tyr Thr Phe Ile Arg Tyr Pro Glu Met Ser Lys		
85	90	95
Glu Met Phe Lys Leu Tyr Ile Pro Gly Ile Asn Ser Lys Val Thr Glu		
100	105	110
Leu Leu Leu Lys Ala Asp Arg Thr Tyr Ser Gln Met Thr Asp Gly Leu		
115	120	125
Arg Asp Leu Trp Ile Asn Val Leu Ser Lys Leu Ala Ser Lys Asn Asp		
130	135	140
Gly Ser Asn Tyr Asp Leu Asn Glu Glu Ile Asn Asn Ile Ser Lys Val		
145	150	155 160
His Thr Thr Tyr Lys Ser Asp Lys Trp Tyr Asn Pro Phe Lys Thr Trp		
165	170	175
Phe Thr Ile Lys Tyr Asp Met Arg Arg Leu Gln Lys Ala Arg Asn Glu		
180	185	190
Ile Thr Phe Asn Val Gly Lys Asp Tyr Asn Leu Leu Glu Asp Gln Lys		
195	200	205
Asn Phe Leu Leu Ile His Pro Glu Leu Val Leu Ile Leu Asp Lys Gln		
210	215	220
Asn Tyr Asn Gly Tyr Leu Ile Thr Pro Glu Leu Val Leu Met Tyr Cys		
225	230	235 240
Asp Val Val Glu Gly Arg Trp Asn Ile Ser Ala Cys Ala Lys Leu Asp		
245	250	255
Pro Lys Leu Gln Ser Met Tyr Gln Lys Gly Asn Asn Leu Trp Glu Val		
260	265	270
Ile Asp Lys Leu Phe Pro Ile Met Gly Glu Lys Thr Phe Asp Val Ile		
275	280	285
Ser Leu Leu Glu Pro Leu Ala Leu Ser Leu Ile Gln Thr His Asp Pro		
290	295	300

SUBSTITUTE SHEET (RULE 26)

- 285 -

Val Lys Gln Leu Arg Gly Ala Phe Leu Asn His Val Leu Ser Glu Met  
305 310 315 320

Glu Leu Ile Phe Glu Ser Arg Glu Ser Ile Lys Glu Phe Leu Ser Val  
325 330 335

Asp Tyr Ile Asp Lys Ile Leu Asp Ile Phe Asn Lys Ser Thr Ile Asp  
340 345 350

Glu Ile Ala Glu Ile Phe Ser Phe Phe Arg Thr Phe Gly His Pro Pro  
355 360 365

Leu Glu Ala Ser Ile Ala Ala Glu Lys Val Arg Lys Tyr Met Tyr Ile  
370 375 380

Gly Lys Gln Leu Lys Phe Asp Thr Ile Asn Lys Cys His Ala Ile Phe  
385 390 395 400

Cys Thr Ile Ile Ile Asn Gly Tyr Arg Glu Arg His Gly Gly Gln Trp  
405 410 415

Pro Pro Val Thr Leu Pro Asp His Ala His Glu Phe Ile Ile Asn Ala  
420 425 430

Tyr Gly Ser Asn Ser Ala Ile Ser Tyr Glu Asn Ala Val Asp Tyr Tyr  
435 440 445

Gln Ser Phe Ile Gly Ile Lys Phe Asn Lys Phe Ile Glu Pro Gln Leu  
450 455 460

Asp Glu Asp Leu Thr Ile Tyr Met Lys Asp Lys Ala Leu Ser Pro Lys  
465 470 475 480

Lys Ser Asn Trp Asp Thr Val Tyr Pro Ala Ser Asn Leu Leu Tyr Arg  
485 490 495

Thr Asn Ala Ser Asn Glu Ser Arg Arg Leu Val Glu Val Phe Ile Ala  
500 505 510

Asp Ser Lys Phe Asp Pro His Gln Ile Leu Asp Tyr Val Glu Ser Gly  
515 520 525

Asp Trp Leu Asp Asp Pro Glu Phe Asn Ile Ser Tyr Ser Leu Lys Glu  
530 535 540

Lys Glu Ile Lys Gln Glu Gly Arg Leu Phe Ala Lys Met Thr Tyr Lys  
545 550 555 560

Met Arg Ala Thr Gln Val Leu Ser Glu Thr Leu Leu Ala Asn Asn Ile  
565 570 575

SUBSTITUTE SHEET (RULE 26)

- 286 -

Gly Lys Phe Phe Gln Glu Asn Gly Met Val Lys Gly Glu Ile Glu Leu  
 580 585 590

Leu Lys Arg Leu Thr Thr Ile Ser Ile Ser Gly Val Pro Arg Tyr Asn  
 595 600 605

Glu Val Tyr Asn Asn Ser Lys Ser His Thr Asp Asp Leu Lys Thr Tyr  
 610 615 620

Asn Lys Ile Ser Asn Leu Asn Leu Ser Ser Asn Gln Lys Ser Lys Lys  
 625 630 635 640

Phe Glu Phe Lys Ser Thr Asp Ile Tyr Asn Asp Gly Tyr Glu Thr Val  
 645 650 655

Ser Cys Phe Leu Thr Thr Asp Leu Lys Lys Tyr Cys Leu Asn Trp Arg  
 660 665 670

Tyr Glu Ser Thr Ala Leu Phe Gly Glu Thr Cys Asn Gln Ile Phe Gly  
 675 680 685

Leu Asn Lys Leu Phe Asn Trp Leu His Pro Arg Leu Glu Gly Ser Thr  
 690 695 700

Ile Tyr Val Gly Asp Pro Tyr Cys Pro Pro Ser Asp Lys Glu His Ile  
 705 710 715 720

Ser Leu Glu Asp His Pro Asp Ser Gly Phe Tyr Val His Asn Pro Arg  
 725 730 735

Gly Gly Ile Glu Gly Phe Cys Gln Lys Leu Trp Thr Leu Ile Ser Ile  
 740 745 750

Ser Ala Ile His Leu Ala Ala Val Arg Ile Gly Val Arg Val Thr Ala  
 755 760 765

Met Val Gln Gly Asp Asn Gln Ala Ile Ala Val Thr Thr Arg Val Pro  
 770 775 780

Asn Asn Tyr Asp Tyr Arg Val Lys Lys Glu Ile Val Tyr Lys Asp Val  
 785 790 795 800

Val Arg Phe Phe Asp Ser Leu Arg Glu Val Met Asp Asp Leu Gly His  
 805 810 815

Glu Leu Lys Leu Asn Glu Thr Ile Ile Ser Ser Lys Met Phe Ile Tyr  
 820 825 830

Ser Lys Arg Ile Tyr Tyr Asp Gly Arg Ile Leu Pro Gln Ala Leu Lys  
 835 840 845

Ala Leu Ser Arg Cys Val Phe Trp Ser Glu Thr Val Ile Asp Glu Thr

SUBSTITUTE SHEET (RULE 26)

- 287 -

850					855					860					
Arg 865	Ser	Ala	Ser	Ser	Asn 870	Leu	Ala	Thr	Ser	Phe 875	Ala	Lys	Ala	Ile	Glu 880
Asn	Gly	Tyr	Ser	Pro 885	Val	Leu	Gly	Tyr	Ala 890	Cys	Ser	Ile	Phe	Lys	Asn 895
Ile	Gln	Gln	Leu	Tyr 900	Ile	Ala	Leu	Gly 905	Met	Asn	Ile	Asn	Pro	Thr	Ile 910
Thr	Gln	Asn	Ile	Arg 915	Asp	Gln	Tyr 920	Phe	Arg	Asn	Pro	Asn	Trp	Met	Gln 925
Tyr	Ala	Ser	Leu	Ile 930	Pro	Ala	Ser 935	Val	Gly	Gly	Phe	Asn	His	Met	Ala 940
Met	Ser	Arg	Cys	Phe 945	Val	Arg	Asn 950	Ile	Gly	Asp	Pro	Ser	Val	Ala	Ala 960
Leu	Ala	Asp	Ile	Lys 965	Arg	Phe	Ile 970	Lys	Ala	Asn	Leu	Leu	Asp	Arg	Ser 975
Val	Leu	Tyr	Arg 980	Ile	Met	Asn	Gln 985	Glu	Pro	Gly	Glu	Ser	Ser	Phe	Phe 990
Asp	Trp	Ala	Ser 995	Asp	Pro	Tyr	Ser 1000	Cys	Asn	Leu	Pro	Gln	Ser	Gln	Asn 1005
Ile	Thr	Thr	Met 1010	Ile	Lys	Asn	Ile 1015	Thr	Ala	Arg	Asn	Val	Leu	Gln	Asp 1020
Ser	Pro	Asn	Pro 1025	Leu	Leu	Ser	Gly 1030	Leu	Phe	Thr	Asn	Thr	Met	Ile	Glu 1040
Glu	Asp	Glu	Glu 1045	Leu	Ala	Glu	Phe 1050	Leu	Met	Asp	Arg	Lys	Val	Ile	Leu 1055
Pro	Arg	Val	Ala 1060	His	Asp	Ile	Leu 1065	Asp	Asn	Ser	Leu	Thr	Gly	Ile	Arg 1070
Asn	Ala	Ile	Ala 1075	Gly	Met	Leu	Asp 1080	Thr	Thr	Lys	Ser	Leu	Ile	Arg	Val 1085
Gly	Ile	Asn	Arg 1090	Gly	Gly	Leu	Thr 1095	Tyr	Ser	Leu	Leu	Arg	Lys	Ile	Ser 1100
Asn	Tyr	Asp	Leu 1105	Val	Gln	Tyr	Glu 1110	Thr	Leu	Ser	Arg	Thr	Leu	Arg	Leu 1120
Ile	Val	Ser	Asp 1125	Lys	Ile	Lys	Tyr 1130	Glu	Asp	Met	Cys	Ser	Val	Asp	Leu 1135

**SUBSTITUTE SHEET (RULE 26)**

- 288 -

Ala Ile Ala Leu Arg Gln Lys Met Trp Ile His Leu Ser Gly Gly Arg  
1140 1145 1150

Met Ile Ser Gly Leu Glu Thr Pro Asp Pro Leu Glu Leu Leu Ser Gly  
1155 1160 1165

Val Val Ile Thr Gly Ser Glu His Cys Lys Ile Cys Tyr Ser Ser Asp  
1170 1175 1180

Gly Thr Asn Pro Tyr Thr Trp Met Tyr Leu Pro Gly Asn Ile Lys Ile  
1185 1190 1195 1200

Gly Ser Ala Glu Thr Gly Ile Ser Ser Leu Arg Val Pro Tyr Phe Gly  
1205 1210 1215

Ser Val Thr Asp Glu Arg Ser Glu Ala Gln Leu Gly Tyr Ile Lys Asn  
1220 1225 1230

Leu Ser Lys Pro Ala Lys Ala Ala Ile Arg Ile Ala Met Ile Tyr Thr  
1235 1240 1245

Trp Ala Phe Gly Asn Asp Glu Ile Ser Trp Met Glu Ala Ser Gln Ile  
1250 1255 1260

Ala Gln Thr Arg Ala Asn Phe Thr Leu Asp Ser Leu Lys Ile Leu Thr  
1265 1270 1275 1280

Pro Val Ala Thr Ser Thr Asn Leu Ser His Arg Phe Lys Asp Thr Ala  
1285 1290 1295

Thr Gln Met Lys Phe Ser Ser Thr Ser Leu Ile Arg Val Ser Arg Phe  
1300 1305 1310

Ile Thr Met Ser Asn Asp Asn Met Ser Ile Lys Glu Ala Asn Glu Thr  
1315 1320 1325

Lys Asp Thr Asn Leu Ile Tyr Gln Gln Ile Met Leu Thr Gly Leu Ser  
1330 1335 1340

Val Phe Glu Tyr Leu Phe Arg Leu Lys Glu Thr Thr Gly His Asn Pro  
1345 1350 1355 1360

Ile Val Met His Leu His Ile Glu Asp Glu Cys Cys Ile Lys Glu Ser  
1365 1370 1375

Phe Asn Asp Glu His Ile Asn Pro Glu Ser Thr Leu Glu Leu Ile Arg  
1380 1385 1390

Tyr Pro Glu Ser Asn Glu Phe Ile Tyr Asp Lys Asp Pro Leu Lys Asp  
1395 1400 1405

SUBSTITUTE SHEET (RULE 26)



- 289 -

Val Asp Leu Ser Lys Leu Met Val Ile Lys Asp His Ser Tyr Thr Ile		
1410	1415	1420
Asp Met Asn Tyr Trp Asp Asp Thr Asp Ile Ile His Ala Ile Ser Ile		
1425	1430	1435 1440
Cys Thr Ala Ile Thr Ile Ala Asp Thr Met Ser Gln Leu Asp Arg Asp		
	1445	1450 1455
Asn Leu Lys Glu Ile Ile Val Ile Ala Asn Asp Asp Asp Ile Asn Ser		
	1460	1465 1470
Leu Ile Thr Glu Phe Leu Thr Leu Asp Ile Leu Val Phe Leu Lys Thr		
	1475	1480 1485
Phe Gly Gly Leu Leu Val Asn Gln Phe Ala Tyr Thr Leu Tyr Ser Leu		
	1490	1495 1500
Lys Ile Glu Gly Arg Asp Leu Ile Trp Asp Tyr Ile Met Arg Thr Leu		
1505	1510	1515 1520
Arg Asp Thr Ser His Ser Ile Leu Lys Val Leu Ser Asn Ala Leu Ser		
	1525	1530 1535
His Pro Lys Val Phe Lys Arg Phe Trp Asp Cys Gly Val Leu Asn Pro		
	1540	1545 1550
Ile Tyr Gly Pro Asn Ile Ala Ser Gln Asp Gln Ile Lys Leu Ala Leu		
	1555	1560 1565
Ser Ile Cys Glu Tyr Ser Leu Asp Leu Phe Met Arg Glu Trp Leu Asn		
	1570	1575 1580
Gly Val Ser Leu Glu Ile Tyr Ile Cys Asp Ser Asp Met Glu Val Ala		
1585	1590	1595 1600
Asn Asp Arg Lys Gln Ala Phe Ile Ser Arg His Leu Ser Phe Val Cys		
	1605	1610 1615
Cys Leu Ala Glu Ile Ala Ser Phe Gly Pro Asn Leu Leu Asn Leu Thr		
	1620	1625 1630
Tyr Leu Glu Arg Leu Asp Leu Leu Lys Gln Tyr Leu Glu Leu Asn Ile		
	1635	1640 1645
Lys Glu Asp Pro Thr Leu Lys Tyr Val Gln Ile Ser Gly Leu Leu Ile		
	1650	1655 1660
Lys Ser Phe Pro Ser Thr Val Thr Tyr Val Arg Lys Thr Ala Ile Lys		
1665	1670	1675 1680
Tyr Leu Arg Ile Arg Gly Ile Ser Pro Pro Glu Val Ile Asp Asp Trp		

SUBSTITUTE SHEET (RULE 26)

- 290 -

1685	1690	1695
Asp Pro Val Glu Asp Glu Asn Met Leu Asp Asn Ile Val Lys Thr Ile 1700	1705	1710
Asn Asp Asn Cys Asn Lys Asp Asn Lys Gly Asn Lys Ile Asn Asn Phe 1715	1720	1725
Trp Gly Leu Ala Leu Lys Asn Tyr Gln Val Leu Lys Ile Arg Ser Ile 1730	1735	1740
Thr Ser Asp Ser Asp Asp Asn Asp Arg Leu Asp Ala Asn Thr Ser Gly 1745	1750	1755 1760
Leu Thr Leu Pro Gln Gly Gly Asn Tyr Leu Ser His Gln Leu Arg Leu 1765	1770	1775
Phe Gly Ile Asn Ser Thr Ser Cys Leu Lys Ala Leu Glu Leu Ser Gln 1780	1785	1790
Ile Leu Met Lys Glu Val Asn Lys Asp Lys Asp Arg Leu Phe Leu Gly 1795	1800	1805
Glu Gly Ala Gly Ala Met Leu Ala Cys Tyr Asp Ala Thr Leu Gly Pro 1810	1815	1820
Ala Val Asn Tyr Tyr Asn Ser Gly Leu Asn Ile Thr Asp Val Ile Gly 1825	1830	1835 1840
Gln Arg Glu Leu Lys Ile Phe Pro Ser Glu Val Ser Leu Val Gly Lys 1845	1850	1855
Lys Leu Gly Asn Val Thr Gln Ile Leu Asn Arg Val Lys Val Leu Phe 1860	1865	1870
Asn Gly Asn Pro Asn Ser Thr Trp Ile Gly Asn Met Glu Cys Glu Ser 1875	1880	1885
Leu Ile Trp Ser Glu Leu Asn Asp Lys Ser Ile Gly Leu Val His Cys 1890	1895	1900
Asp Met Glu Gly Ala Ile Gly Lys Ser Glu Glu Thr Val Leu His Glu 1905	1910	1915 1920
His Tyr Ser Val Ile Arg Ile Thr Tyr Leu Ile Gly Asp Asp Asp Val 1925	1930	1935
Val Leu Val Ser Lys Ile Ile Pro Thr Ile Thr Pro Asn Trp Ser Arg 1940	1945	1950
Ile Leu Tyr Leu Tyr Lys Leu Tyr Trp Lys Asp Val Ser Ile Ile Ser 1955	1960	1965

SUBSTITUTE SHEET (RULE 26)

- 291 -

Leu Lys Thr Ser Asn Pro Ala Ser Thr Glu Leu Tyr Leu Ile Ser Lys  
 1970 1975 1980

Asp Ala Tyr Cys Thr Ile Met Glu Pro Ser Glu Ile Val Leu Ser Lys  
 1985 1990 1995 2000

Leu Lys Arg Leu Ser Leu Leu Glu Glu Asn Asn Leu Leu Lys Trp Ile  
 2005 2010 2015

Ile Leu Ser Lys Lys Arg Asn Asn Glu Trp Leu His His Glu Ile Lys  
 2020 2025 2030

Glu Gly Glu Arg Asp Tyr Gly Ile Met Arg Pro Tyr His Met Ala Leu  
 2035 2040 2045

Gln Ile Phe Gly Phe Gln Ile Asn Leu Asn His Leu Ala Lys Glu Phe  
 2050 2055 2060

Leu Ser Thr Pro Asp Leu Thr Asn Ile Asn Asn Ile Ile Gln Ser Phe  
 2065 2070 2075 2080

Gln Arg Thr Ile Lys Asp Val Leu Phe Glu Trp Ile Asn Ile Thr His  
 2085 2090 2095

Asp Asp Lys Arg His Lys Leu Gly Gly Arg Tyr Asn Ile Phe Pro Leu  
 2100 2105 2110

Lys Asn Lys Gly Lys Leu Arg Leu Leu Ser Arg Arg Leu Val Leu Ser  
 2115 2120 2125

Trp Ile Ser Leu Ser Leu Ser Thr Arg Leu Leu Thr Gly Arg Phe Pro  
 2130 2135 2140

Asp Glu Lys Phe Glu His Arg Ala Gln Thr Gly Tyr Val Ser Leu Ala  
 2145 2150 2155 2160

Asp Thr Asp Leu Glu Ser Leu Lys Leu Leu Ser Lys Asn Ile Ile Lys  
 2165 2170 2175

Asn Tyr Arg Glu Cys Ile Gly Ser Ile Ser Tyr Trp Phe Leu Thr Lys  
 2180 2185 2190

Glu Val Lys Ile Leu Met Lys Leu Ile Gly Gly Ala Lys Leu Leu Gly  
 2195 2200 2205

Ile Pro Arg Gln Tyr Lys Glu Pro Glu Asp Gln Leu Leu Glu Asn Tyr  
 2210 2215 2220

Asn Gln His Asp Glu Phe Asp Ile Asp  
 2225 2230

SUBSTITUTE SHEET (RULE 26)

- 292 -

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15218 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACGCGAAAAA ATGCGTACTA CAACTTGCA CATTCGAAAA AAATGGGGCA AATAAGAACT	60
TGATAAGTGC TATTTAAGTC TAACCTTTTC AATCAGAAAT GGGGTGCAAT TCACTGAGCA	120
TGATAAAGGT TAGATTACAA AATTTATTTG ACAATGACGA AGTAGCATTG TTAAAAATAA	180
CATGTTATAC TGATAAATTA ATTCTTCTGA CCAATGCATT AGCCAAAGCA GCAATACATA	240
CAATTAAATT AAACGGCATA GTTTTATAC ATGTTATAAC AAGCAGTGAA GTGTGCCCTG	300
ATAACAATAT TGTAGTGAAA TCTAACTTTA CAACAATGCC AATACTACAA AATGGAGGAT	360
ACATATGGGA ATTGATTGAG TTGACACACT GCTCTCAATT AAACGGTTTA ATGGATGATA	420
ATTGTGAAAT CAAATTTTCT AAAAGACTAA GTGACTCAGT AATGACTAAT TATATGAATC	480
AAATATCTGA CTTACTTGGG CTTGATCTCA ATTCATGAAT TATGTTTAGT CTAATTCAAT	540
AGACATGTGT TTATTACCAT TTTAGTTAAT ATAAAACTC ATCAAAGGGA AATGGGGCAA	600
ATAAACTCAC CTAATCAATC AAACCATGAG CACTACAAAT GACAACACTA CTATGCAAAG	660
ATTGATGATC ACAGACATGA GACCCCTGTC AATGGATTCA ATAATAACAT CTCTTACCAA	720
AGAAATCATC ACACACAAAT TCATATACTT GATAAACAAT GAATGTATTG TAAGAAAAC	780
TGATGAAAGA CAAGCTACAT TTACATTCTT AGTCAATTAT GAGATGAAGC TACTGCACAA	840
AGTAGGGAGT ACCAAATACA AAAAATACAC TGAATATAAT ACAAATATG GCACTTTCCC	900
CATGCCTATA TTTATCAATC ACGGCGGGTT TCTAGAATGT ATTGGCATTG AGCCTACAA	960
ACACACTCCT ATAATATACA AATATGACCT CAACCCGTGA ATTCCAACAA AAAAACCAAC	1020
CCAACCAAC CAACTATTC CTCAAACAAC AGTGCTCAAT AGTTAAGAAG GAGCTAATCC	1080
ATTTTAGTAA TTAAAAATAA AAGTAAAGCC AATAACATAA ATTGGGGCAA ATACAAAGAT	1140

SUBSTITUTE SHEET (RULE 26)

- 293 -

GGCTCTTAGC AAAGTCAAGT TGAATGATAC ATTAATAAG GATCAGCTGC TGTCATCCAG 1200  
CAAATACACT ATTCAACGTA GTACAGGAGA TAATATTGAC ACTCCCAATT ATGATGTGCA 1260  
AAAACACCTA AACAACTAT GTGGTATGCT ATTAATCACT GAAGATGCAA ATCATAAATT 1320  
CACAGGATTA ATAGGTATGT TATATGCTAT GTCCAGGTTA GGAAGGGAAG ACACTATAAA 1380  
GATACTTAAA GATGCTGGAT ATCATGTTAA AGCTAATGGA GTAGATATAA CAACATATCG 1440  
TCAAGATATA AATGGAAAGG AAATGAAATT CGAAGTATTA ACATTATCAA GCTTGACATC 1500  
AGAAATACAA GTCAATATTG AGATAGAATC TAGAAAGTCC TACAAAAAAA TGCTAAAAGA 1560  
GATGGGAGAA GTGGCTCCAG AATATAGGCA TGATTCTCCA GACTGTGGGA TGATAATACT 1620  
GTGTATAGCT GCACTTGTGA TAACCAAATT AGCAGCAGGA GACAGATCAG GTCTTACAGC 1680  
AGTAATTAGG AGGGCAAACA ATGTCTTAAA AAACGAAATA AAACGATACA AGGGCCTCAT 1740  
ACCAAAGGAT ATAGCTAACA GTTTTATGA AGTGTGTTGAA AAACACCCTC ATCTTATAGA 1800  
TGTTTTCGTG CACTTTGGCA TTGCACAATC ATCCACAAGA GGGGGTAGTA GAGTTGAAGG 1860  
AATCTTTGCA GGATTGTTTA TGAATGCCTA TGGTTCAGGG CAAGTAATGC TAAGATGGGG 1920  
AGTTTTAGCC AAATCTGTAA AAAATATCAT GCTAGGACAT GCTAGTGTCC AGGCAGAAAT 1980  
GGAGCAAGTT GTGGAAGTCT ATGAGTATGC ACAGAAGTTG GGAGGAGAAG CTGGATTCTA 2040  
CCATATATTG AACAATCCAA AAGCATCATT GCTGTCATTA ACTCAATTTC CCAACTTCTC 2100  
AAGTGTGGTC CTAGGCAATG CAGCAGGTCT AGGCATAATG GGAGAGTATA GAGGTACACC 2160  
AAGAAACCAG GATCTTTATG ATGCAGCTAA AGCATATGCA GAGCAACTCA AAGAAAATGG 2220  
AGTAATAAAC TACAGTGTAT TAGACTTAAC AGCAGAAGAA TTGGAAGCCA TAAAGCATCA 2280  
ACTCAACCCC AAAGAAGATG ATGTAGAGCT TTAAGTTAAC AAAAAATACG GGGCAAATAA 2340  
GTCAACATGG AGAAGTTTGC ACCTGAATTT CATGGAGAAG ATGCAAATAA CAAAGCTACC 2400  
AAATTCCTAG AATCAATAAA GGGCAAGTTC GCATCATCCA AAGATCCTAA GAAGAAAGAT 2460  
AGCATAATAT CTGTAACTC AATAGATATA GAAGTAACTA AAGAGAGCCC GATAACATCT 2520  
GGCACCAACA TCATCAATCC AACAAGTGAA GCCGACAGTA CCCCAGAAAC AAAAGCCAAC 2580  
TACCCAAGAA AACCCCTAGT AAGCTTCAA GAAGATCTCA CCCCAGTGA CAACCCTTTT 2640  
TCTAAGTTGT ACAAGGAAAC AATAGAAACA TTTGATAACA ATGAAGAAGA ATCTAGCTAC 2700

SUBSTITUTE SHEET (RULE 26)

- 294 -

TCATATGAAG AGATAAATGA TCAAACAAAT GACAACATTA CAGCAAGACT AGATAGAATT 2760  
GATGAAAAAT TAAGTGAAAT ATTAGGAATG CTCCATACAT TAGTAGTTGC AAGTGCAGGA 2820  
CCCACCTCAG CTCGCGATGG AATAAGAGAT GCTATGGTTG GTCTAAGAGA AGAGATGATA 2880  
GAAAAAATAA GAGCGGAAGC ATTAATGACC AATGATAGGT TAGAGGCTAT GGCAAGACTT 2940  
AGGAATGAGG AAAGCGAAAA AATGGCAAAA GACACCTCAG ATGAAGTGTC TCTTAATCCA 3000  
ACTTCCAAAA AATTGAGTGA CTTGTTGGAA GACAACGATA GTGACAATGA TCTATCACTT 3060  
GATGATTTTT GATCAGCGAT CAACTCACTC AGCAATCAAC AACATCAATA AAACAGACAT 3120  
CAATCCATTG AATCAACTGC CAGACCGAAC AAACAAACGT CCATCAGTAG AACCACCAAC 3180  
CAATCAATCA ACCAATTGAT CAATCAGCAA CCCGACAAAA TTAACAATAT AGTAACAAAA 3240  
AAAGAACAAG ATGGGGCAAA TATGGAAACA TACGTGAACA AGCTTCACGA AGGCTCCACA 3300  
TACACAGCAG CTGTTCACTA CAATGTTCTA GAAAAAGATG ATGATCCTGC ATCACTAACA 3360  
ATATGGGTGC CTATGTTCCA GTCATCTGTG CCAGCAGACT TGCTCATAAA AGAAGTTGCA 3420  
AGCATCAATA TACTAGTGAA GCAGATCTCT ACGCCCAAAG GACCTTCACT ACGAGTCACG 3480  
ATTAAGTCAA GAAGTGCTGT GCTGGCTCAA ATGCCTAGTA ATTTTCATCAT AAGCGCAAAT 3540  
GTATCATTAG ATGAAAGAAG CAAATTAGCA TATGATGTAA CTACACCTTG TGAATCAAA 3600  
GCATGCAGTC TAACATGCTT AAAAGTAAAA AGTATGTTAA CTACAGTCAA AGATCTTACC 3660  
ATGAAGACAT TCAACCCAC TCATGAGATC ATTGCTCTAT GTGAATTGA AAATATTATG 3720  
ACATCAAAAA GAGTAATAAT ACCAACCTAT CTAAGATCAA TTAGTGTCAA GAACAAGGAT 3780  
CTGAACTCAC TAGAAAAATAT AGCAACCACC GAATTCAAAA ATGCTATCAC CAATGCAAAA 3840  
ATTATTCCTT ATGCAGGATT AGTGTTAGTT ATCACAGTTA CTGACAATAA AGGAGCATTC 3900  
AAATATATCA AACCACAGAG TCAATTTATA GTAGATCTTG GTGCCTACCT AGAAAAAGAG 3960  
AGCATATATT ATGTGACTAC TAATTGGAAG CATAAGCTA CACGTTTTTC AATCAAACCA 4020  
CTAGAGGATT AACTTAAAT ATCAACACTG AATGACAGGT CCACATATAT CCTCAAACCTA 4080  
CACACTATAT CCAAACATCA TAAACATCTA CACTACACAC TTCATCACAC AAACCAATCC 4140  
CACTCAAAAT CCAAAATCAC TACCAGCCAC TATCTGCTAG ACCTAGAGTG CGAATAGGTA 4200  
AATAAAACCA AAATATGGGG TAAATAGACA TTAGTTAGAG TTCAATCAAT CTTAACAACC 4260

SUBSTITUTE SHEET (RULE 26)

- 295 -

ATTATACCG CCAATTCAAC ACATATACTA TAAATCTTAA AATGGGAAAT ACATCCATCA 4320  
CAATAGAATT CACAAGCAAA TTTTGGCCCT ATTTTACACT AATACATATG ATCTTAACTC 4380  
TAATCTTTTT ACTAATTATA ATCACTATTA TGATTGCAAT ACTAAATAAG CTAAGTGAAC 4440  
ATAAAGCATT CTGTAACAAA ACTCTTGAAC TAGGACAGAT GTATCAAATC AACACATAGA 4500  
GTTCTACCAT TATGCTGTGT CAAATTATAA TCCTGTATAT ATAAACAAAC AAATCCAATC 4560  
TTCTCACAGA GTCATGGTGT CGCAAAACCA CGCTAACTAT CATGGTAGCA TAGAGTAGTT 4620  
ATTTAAAAAT TAACATAATG ATGAATTGTT AGTATGAGAT CAAAAACAAC ATTGGGGCAA 4680  
ATGCAACCAT GTCCAAACAC AAGAATCAAC GCACTGCCAG GACTCTAGAA AAGACCTGGG 4740  
ATACTCTTAA TCATCTAATT GTAATATCCT CTTGTTTATA CAGATTAAAT TTAAATCTA 4800  
TAGCACAAAT AGCACTATCA GTTTTGGCAA TGATAATCTC AACCTCTCTC ATAATTGCAG 4860  
CCATAATATT CATCATCTCT GCCAATCACA AAGTTACACT AACAACGGTC ACAGTTCAAA 4920  
CAATAAAAAA CCACACTGAA AAAAACATCA CCACCTACCC TACTCAAGTC TCACCAGAAA 4980  
GGGTTAGTTC ATCCAAGCAA CCCACAACCA CATCACCAAT CCACACAAGT TCAGCTACAA 5040  
CATCACCCAA TACAAAATCA GAAACACACC ATACAACAGC ACAAACCAAA GGCAGAACCA 5100  
CCACTTCAAC ACAGACCAAC AAGCCAAGCA CAAAACCACG TCCAAAAAAT CCACCAAAAA 5160  
AAGATGATTA CCATTTTGAA GTGTTCAACT TCGTTCCTG CAGTATATGT GGCAACAATC 5220  
AACTTTGCAA ATCCATCTGC AAAACAATAC CAAGCAACAA ACCAAAGAAG AAACCAACCA 5280  
TCAAACCCAC AAACAAACCA ACCACCAAAA CCACAAACAA AAGAGACCCA AAAACACCAG 5340  
CCAAAACGAC GAAAAAAGAA ACTACCACCA ACCCAACAAA AAACTAACC CTCAAGACCA 5400  
CAGAAAGAGA CACCAGCACC TCACAATCCA CTGCACTCGA CACAACCACA TTAAACACA 5460  
CAGTCCAACA GCAATCCCTC CTCTCAACCA CCCCCGAAA CACACCCAAC TCCACACAAA 5520  
CACCCACAGC ATCCGAGCCC TCCACACCA ACTCCACCCA AAAAACCAG CCACATGCTT 5580  
AGTTATTCAA AACTACATC TTAGCAGAGA ACCGTGATCT ATCAAGCAAG AACGAAATTA 5640  
AACTGGGGC AAATAACCAT GGAGTTGATG ATCCACAAGT CAAGTGCAAT CTTCTTAATC 5700  
CTTGCTATTA ATGCATTGTA CCTCACCTCA AGTCAGAACA TAACTGAGGA GTTTTACCAA 5760  
TCGACATGTA GTGCAGTTAG CAGAGGTTAT TTTAGTGCTT TAAGAACAGG TTGGTATACT 5820

SUBSTITUTE SHEET (RULE 26)

- 296 -

AGTGTCATAA CAATAGAATT AAGTAATATA AAAGAAACCA AATGCAATGG AACTGACACT 5880  
AAAGTAAAC TTATGAAACA AGAATTAGAT AAGTATAAGA ATGCAGTAAC AGAATTACAG 5940  
CTACTTATGC AAAACACACC AGCTGTCAAC AACCGGGCCA GAAGAGAAGC ACCACAGTAT 6000  
ATGAACTACA CAATCAATAC CACTAAAAAC CTAAATGTAT CAATAAGCAA GAAGAGGAAA 6060  
CGAAGATTTT TAGGCTTCTT GTTAGGTGTG GGATCTGCAA TAGCAAGTGG TATAGCTGTA 6120  
TCAAAAGTTC TACACCTTGA AGGAGAAGTG AACAAAGATCA AAAATGCTTT GTTGTCTACA 6180  
AACAAAGCTG TAGTCAGTTT ATCAAATGGG GTCAGTGTTT TAACCAGCAA AGTGTTAGAT 6240  
CTCAAGAATT ACATAAATAA CCAATTATTA CCCATAGTAA ATCAACAGAG CTGTCGCATC 6300  
TCCAACATTG AACAGTTAT AGAATTCCAG CAGAAGAACA GCAGATTGTT GGAAATCACC 6360  
AGAGAATTTA GTGTCAATGC AGGTGTAACA ACACCTTTAA GCACTTACAT GTTGACAAAC 6420  
AGTGAGTTAC TATCATTAAAT CAATGATATG CCTATAACAA ATGATCAGAA AAAATTAATG 6480  
TCAAGCAATG TTCAGATAGT AAGGCAACAA AGTTATTCCA TCATGTCTAT AATAAAGGAA 6540  
GAAGTCCTTG CATATGTTGT ACAGCTGCCT ATCTATGGTG TAATAGATAC ACCTTGCTGG 6600  
AAATTGCACA CATCGCCTCT ATGCACTACC AACATCAAAG AAGGATCAAA TATTTGTTTA 6660  
ACAAGGACTG ATAGAGGATG GTATTGTGAT AATGCAGGAT CAGTATCCTT CTTCCACAG 6720  
GCTGACACTT GTAAAGTACA GTCCAATCGA GTATTTGTG ACACATGAA CAGTTTGACA 6780  
TTACCAAGTG AAGTCAGCCT TTGTAACACT GACATATTCA ATTCCAAGTA TGA CTGCAAA 6840  
ATTATGACAT CAAAAACAGA CATAAGCAGC TCAGTAATTA CTTCTCTTGG AGCTATAGTG 6900  
TCATGCTATG GTAAACTAA ATGCACTGCA TCCAACAAA ATCGTGGGAT TATAAAGACA 6960  
TTTTCTAATG GTTGTGACTA TGTGTCAAAC AAAGGAGTAG ATACTGTGTC AGTGGGCAAC 7020  
ACTTTATACT ATGTAAACAA GCTGGAAGGC AAGAACCTTT ATGTAAAAGG GGAACCTATA 7080  
ATAAATTACT ATGACCCCTCT AGTGTTTCCT TCTGATGAGT TTGATGCATC AATATCTCAA 7140  
GTCAATGAAA AAATCAATCA AAGTTTAGCT TTTATTCGTA GATCTGATGA ATTACTACAT 7200  
AATGTAAATA CTGGCAAATC TACTACAAAT ATTATGATAA CTACAATTAT TATAGTAATC 7260  
ATTGTAGTAT TGTATCATT AATAGCTATT GGT TACTGT TGTATTGTAA AGCCAAAAAC 7320  
ACACCAGTTA CACTAAGCAA AGACCAACTA AGTGGAATCA ATAATATTGC ATTCAGCAAA 7380

SUBSTITUTE SHEET (RULE 26)



- 297 -

TAGACAAAA ACCACCTGAT CATGTTTCAA CAACAATCTG CTGACCACCA ATCCCAAATC 7440  
AACTTACAAC AAATATTCA ACATCACAGT ACAGGCTGAA TCATTTCTC ACATCATGCT 7500  
ACCCACATAA CTAAGCTAGA TCCTTAACTT ATAGTTACAT AAAAACCTCA AGTATCACAA 7560  
TCAACCACTA AATCAACACA TCATTACAA AATTAACAGC TGGGGCAAAT ATGTCGCGAA 7620  
GAAATCCTTG TAAATTTGAG ATTAGAGGTC ATTGCTTGAA TGGTAGAAGA TGCTACTACA 7680  
GTCATAATTA CTTTGAATGG CCTCCTCATG CATTACTAGT GAGGCAAAAC TTCATGTTAA 7740  
ACAAGATACT CAAGTCAATG GACAAAAGCA TAGACACTTT GTCTGAAATA AGTGGAGCTG 7800  
CTGAAGTGGG TAGAACAGAA GAATATGCTC TTGGTATAGT TGGAGTGCTA GAGAGTTACA 7860  
TAGGATCTAT AAACAACATA ACAAACAAT CAGCATGTGT TGCTATGAGT AAACCTCTTA 7920  
TTGAGATCAA TAGTGATGAC ATTAAAAAGC TTAGAGATAA TGAAGAACCC AATTCACCTA 7980  
AGATAAGAGT GTACAATACT GTTATATCAT ACATTGAGAG CAATAGAAAA AACAACAAGC 8040  
AAACCATCCA TCTGCTCAAG AGACTACCAG CAGACGTGCT GAAGAAGACA ATAAAGAACA 8100  
CATTAGATAT CCACAAAAGC ATAACCATAA GCAATCCAAA AGAGTCAACT GTGAATGATC 8160  
AAAATGACCA AACCAAAAT AATGATATTA CCGATAAAT ATCCTTGTA TATATCATCC 8220  
ATATTGATCT CAAGTGAAAG CATGGTTGCT ACATTCAATC ATAAAAACAT ATTACAATTT 8280  
AACCATAACT ATTTGGATAA CCACCAGCGT TTATTAAATC ATATATTTGA TGAAATTCAT 8340  
TGGACACCTA AAAACTTATT AGATGCCACT CAACAATTTT TCCAACATCT TAACATCCCT 8400  
GAAGATATAT ATACAGTATA TATATTAGTG TCATAATGCT TGACCATAAC GACTCTATGT 8460  
CATCCAACCA TAAACTATT TTGATAAGGT TATGGGACAA AATGGATCCC ATTATTAATG 8520  
GAAACTCTGC TAATGTGTAT CTAAGTATA GTTATTTAAA AGGTGTTATC TCTTTTTCAG 8580  
AGTGTAATGC TTTAGGGAGT TATCTTTTAA ACGGCCCTTA TCTTAAAAAT GATTACACCA 8640  
ACTTAATTAG TAGACAAAGC CCACTACTAG AGCATATGAA TCTTAAAAAA CTAAGTATAA 8700  
CACAGTCATT AATATCTAGA TATCATAAAG GTGAAGTAA ATTAGAAGAA CCAACTTATT 8760  
TCCAGTCATT ACTTATGACA TATAAAAGTA TGTCCTCGTC TGAACAAAT GCTACAACTA 8820  
ACTTACTTAA AAAATAATA CGAAGAGCCA TAGAAATAAG TGATGTAAAG GTGTACGCCA 8880  
TCTTGAATAA ACTAGGATTA AAGGAAAAGG ACAGAGTTAA GCCCAACAAT AATTCAGGTG 8940

SUBSTITUTE SHEET (RULE 26)

- 298 -

ATGAAAAC TC AGTACTTACA ACCATAATTA AAGATGATAT ACTTTCGGCT GTGGAAAACA 9000  
ATCAATCATA TACAAATTCA GACAAAAGTC ACTCAGTAAA TCAAAATATC ACTATCAAAA 9060  
CAACACTCTT GAAAAAATTG ATGTGTTCAA TGCAACATCC TCCATCATGG TTAATACACT 9120  
GGTTC AATTT ATATACAAA TTAAATAACA TATTAACACA ATATCGATCA AATGAGGTAA 9180  
AAAGTCATGG GTTTATATTA ATAGATAATC AAAC TTTAAG TGGTTTTCAG TTTATTTTAA 9240  
ATCAATATGG TTGTATCGTT TATCATAAAG GACTCAAAAA AATCACAAC TACTACTTACA 9300  
ATCAATTTTT GACATGGAAA GACATCAGCC TTAGCAGATT AAATGTTTGC TTAATTACTT 9360  
GGATAAGTAA TTGTTTAAAT ACATTAAACA AAAGCTTAGG GCTGAGATGT GGATTCAATA 9420  
ATGTTGTGTT ATCACAATTA TTTCTTTATG GAGATTGTAT ACTGAAATTA TTTCATAATG 9480  
AAGGCTTCTA CATAATAAAA GAAGTAGAGG GATT TATTAT GTCTTTAATT CTAACATAA 9540  
CAGAAGAAGA TCAATTTAGG AAACGATTTT ATAATAGCAT GCTAAATAAC ATCAGATG 9600  
CAGCTATTAA GGCTCAAAAG GACCTACTAT CAAGAGTATG TCACACTTTA TTAGACAAGA 9660  
CAGTGCTGA TAATATCATA AATGGTAAAT GGATAATCCT ATTAAGTAAA TTTCTTAAAT 9720  
TGATTAAGCT TGCAGGTGAT AATAATCTCA ATAAC TTGAG TGA GCTATAT TTTCTCTTCA 9780  
GAATCTTTGG ACATCCAATG GTCGATGAAA GACAAGCAAT GGATTCTGTA AGAATTAAC T 9840  
GTAATGAAAC TAAGTTCTAC TTATTAAGTA GTCTAAGTAC ATTAAGAGGT GCTTTCATTT 9900  
ATAGAATCAT AAAAGGGTTT GTAAATACCT ACAACAGATG GCCCACCTTA AGGAATGCTA 9960  
TTGTCCTACC TCTAAGATGG TTAACTACT ATAACTTAA TACTTATCCA TCTCTACTTG 10020  
AAATCACAGA AAATGATTTG ATTATTTTAT CAGGATTGCG GTTCTATCGT GAGTTTCATC 10080  
TGCCTAAAAA AGTGGATCTT GAAATGATAA TAAATGACAA AGCCATTTCA CCTCCAAAAG 10140  
ATCTAATATG GACTAGTTTT CCTAGAAATT ACATGCCATC ACATATACAA AATTATATAG 10200  
AACATGAAAA GTTGAAGTTC TCTGAAAGCG ACAGATCGAG AAGAGTACTA GAGTATTACT 10260  
TGAGAGATAA TAAATTCAAT GAATGCGATC TATACAATTG TG TAGTCAAT CAAAGCTATC 10320  
TCAACAACTC TAATCACGTG GTATCACTAA CTGGTAAAGA AAGAGAGCTC AGTGTAGGTA 10380  
GAATGTTTGC TATGCAACCA GGTATGTTTA GGCAATCCA AATCTTAGCA GAGAAAATGA 10440  
TAGCTGAAAA TATTTTACAA TTCTTCCCTG AGAGTTTGAC AAGATATGGT GATCTAGAGC 10500

SUBSTITUTE SHEET (RULE 26)

- 299 -

TTCAAAAGAT ATTAGAATTA AAAGCAGGAA TAAGCAACAA GTCAAATCGT TATAATGATA 10560  
ACTACAACAA TTATATCAGT AAATGTTCTA TCATTACAGA TCTTAGCAAA TTCAATCAGG 10620  
CATTTAGATA TGAAACATCA TGTATCTGCA GTGATGTATT AGATGAACTG CATGGAGTAC 10680  
AATCTCTGTT CTCTTGGTTG CATTTAACAA TACCTCTTGT CACAATAATA TGTACATATA 10740  
GACATGCACC TCCTTTCATA AAGGATCATG TTGTTAATCT TAATGAGGTT GATGAACAAA 10800  
GTGGATTATA CAGATATCAT ATGGGTGGTA TTGAGGGCTG GTGTCAAAAA CTGTGGACCA 10860  
TTGAAGCTAT ATCATTATTA GATCTAATAT CTCTCAAAGG GAAATTCTCT ATCAGAGCTC 10920  
TGATAAATGG TGATAATCAG TCAATTGATA TAAGCAAACC AGTTAGACTT ATAGAGGGTC 10980  
AGACCCATGC ACAAGCAGAT TATTTGTTAG CATTAAATAG CCTTAAATTG TTATATAAAG 11040  
AGTATGCAGG TATAGGCCAT AAGCTTAAGG GAACAGAGAC CTATATATCC CGAGATATGC 11100  
AGTTCATGAG CAAAACAATC CAGCACAATG GAGTGTACTA TCCAGCCAGT ATCAAAAAAG 11160  
TCCTGAGAGT AGGTCCATGG ATAAACACGA TACTTGATGA TTTTAAAGTT AGTTTAGAAT 11220  
CTATAGGCAG CTTAACACAG GAGTTAGAAT ACAGAGGAGA AAGCTTATTA TGCAGTTTAA 11280  
TATTTAGGAA CATTGTTTGA TACAATCAAA TTGCTTTGCA ACTCCGAAAT CATGCATTAT 11340  
GTAACAATAA GCTATATTTA GATATATTGA AAGTATTAAA AACTTAAAA ACTTTTTTTA 11400  
ATCTTGATAG CATTGATATG GCTTTATCAT TGTATATGAA TTTGCCTATG CTGTTTGGTG 11460  
GTGGTGATCC TAATTTGTTA TATCGAAGCT TTTATAGGAG AACTCCAGAC TTCCTTACAG 11520  
AAGCTATAGT ACATTCAGTG TTTGTGTTGA GCTATTATAC TGGTCACGAT TTACAAGATA 11580  
AGCTCCAGGA TCTTCCAGAT GATAGACTGA ACAAATTCTT GACATGTGTC ATCACATTTG 11640  
ATAAAAATCC CAATGCCGAG TTTGTAACAT TGATGAGGGA TCCACAGGCT TTAGGGTCTG 11700  
AAAGGCAAGC TAAAATTACT AGTGAGATTA ATAGATTAGC AGTAACAGAA GTCTTAAGTA 11760  
TAGCCCCAAA CAAAATATTT TCTAAAAGTG CACAACATTA TACTACCACT GAGATTGATC 11820  
TAAATGACAT TATGCAAAAT ATAGAACCAA CTTACCCTCA TGGATTAAGA GTTGTTTATG 11880  
AAAGTTTACC TTTTATAAAA GCAGAAAAAA TAGTTAATCT TATATCAGGA ACAAATCCA 11940  
TAACTAATAT ACTTGAAAAA ACATCAGCAA TAGATACAAC TGATATTAAT AGGGCTACTG 12000  
ATATGATGAG GAAAAATATA ACTTTACTTA TAAGGATACT TCCACTAGAT TGTAACAAAG 12060

SUBSTITUTE SHEET (RULE 26)

- 300 -

ACAAAAGAGA GTTATTAAGT TTAGAAAATC TTAGTATAAC TGAATTAAGC AAGTATGTAA 12120  
GAGAAAGATC TTGGTCATTA TCCAATATAG TAGGAGTAAC ATCGCCAAGT ATTATGTTCA 12180  
CAATGGACAT TAAATATACA ACTAGCACTA TAGCCAGTGG TATAATAATA GAAAAATATA 12240  
ATGTTAATAG TTTAACTCGT GGTGAAAGAG GACCCACCAA GCCATGGGTA GGCTCATCCA 12300  
CGCAGGAGAA AAAACAATG CCAGTGTACA ACAGACAAGT TTTAACC AAA AAGCAAAGAG 12360  
ACCAAATAGA TTTATTAGCA AAATTAGACT GGGTATATGC ATCCATAGAC AACAAAGATG 12420  
AATTCATGGA AGAACTGAGT ACTGGAACAC TTGGACTGTC ATATGAAAAA GCCAAAAAGT 12480  
TGTTTCCACA ATATCTAAGT GTCAATTATT TACACCGTTT AACAGTCAGT AGTAGACCAT 12540  
GTGAATTCCC TGCATCAATA CCAGCTTATA GAACAACAAA TTATCATTTT GATACTAGTC 12600  
CTATCAATCA TGTATTAACA GAAAAGTATG GAGATGAAGA TATCGACATT GTGTTTCAAA 12660  
ATTGCATAAG TTTTGGTCTT AGCCTGATGT CGGTTGTGGA ACAATTCACA AACATATGTC 12720  
CTAATAGAAT TATTCTCATA CCGAAGCTGA ATGAGATACA TTTGATGAAA CCTCCTATAT 12780  
TTACAGGAGA TGTTGATATC ATCAAGTTGA AGCAAGTGAT ACAAAGCAG CACATGTTCC 12840  
TACCAGATAA AATAAGTTTA ACCCAATATG TAGAATTATT CTTAAGTAAC AAAGCACTTA 12900  
AATCTGGATC TCACATCAAC TCTAATTTAA TATTAGTACA TAAAATGTCT GATTATTTTC 12960  
ATAATGCTTA TATTTTAAGT ACTAATTTAG CTGGACATTG GATTCTGATT ATTCACTTA 13020  
TGAAAGATTG AAAAGGTATT TTTGAAAAAG ATTGGGGAGA GGGGTACATA ACTGATCATA 13080  
TGTTTCATTAA TTTGAATGTT TTCTTTAATG CTTATAAGAC TTATTTGCTA TGTTTTCATA 13140  
AAGGTTATGG TAAAGCAAAA TTAGAATGTG ATATGAACAC TTCAGATCTT CTTTGTGTTT 13200  
TGGAGTTAAT AGACAGTAGC TACTGGAAAT CTATGTCTAA AGTTTTCTTA GAACAAAAAG 13260  
TCATAAAATA CATAGTCAAT CAAGACACAA GTTTCGCTAG AATAAAAGGC TGTCACAGTT 13320  
TTAAGTTGTG GTTTTTAAAA CGCCTTAATA ATGCTAAATT TACCGTATGC CCTTGGGTTG 13380  
TTAACATAGA TTATCACCCA ACACACATGA AAGCTATATT ATCTTACATA GATTTAGTTA 13440  
GAATGGGGTT AATAAATGTA GATAAATTAA CCATTAAAAA TAAAAACAAA TTCAATGATG 13500  
AATTTTACAC ATCAAATCTC TTTTACATTA GTTATAACTT TTCAGACAAC ACTCATTTGC 13560  
TAACAAAACA AATAAGAATT GCTAATTCAG AATTAGAAGA TAATTATAAC AAACATATATC 13620

SUBSTITUTE SHEET (RULE 26)

- 301 -

ACCCAACCCC AGAACTTTA GAAAATATGT CATTAAATTCC TGTAAAAAGT AATAATAGTA	13680
ACAAACCTAA ATTTTGTATA AGTGGAATA CCGAATCTAT GATGATGTCA ACATTCTCTA	13740
GTAAATGCA TATTAAATCT TCCACTGTTA CCACAAGATT CAATTATAGC AAACAAGACT	13800
TGTACAATTT ATTTCCAATT GTTGTGATAG ACAAGATTAT AGATCATTCA GGTAATACAG	13860
CAAAATCTAA CCAACTTTAC ACCACCACTT CACATCAGAC ATCTTTAGTA AGGAATAGTG	13920
CATCACTTTA TTGCATGCTT CCTTGGCATC ATGTCAATAG ATTTAACTTT GTATTTAGTT	13980
CCACAGGATG CAAGATCAGT ATAGAGTATA TTTTAAAAGA TCTTAAGATT AAGGACCCCA	14040
GTTGTATAGC ATTCATAGGT GAAGGAGCTG GTAACCTTATT ATTACGTACG GTAGTAGAAC	14100
TTCATCCAGA CATAAGATAC ATTTACAGAA GTTTAAAAGA TTGCAATGAT CATAGTTTAC	14160
CTATTGAATT TCTAAGGTTA TACAACGGGC ATATAAACAT AGATTATGGT GAGAATTTAA	14220
CCATTCCTGC TACAGATGCA ACTAATAACA TTCATTGGTC TTATTTACAT ATAAAATTG	14280
CAGAACCTAT TAGCATCTTT GTCTGCGATG CTGAATTACC TGTTACAGCC AATTGGAGTA	14340
AAATTATAAT TGAATGGAGT AAGCATGTAA GAAAGTGCAA GTACTGTTCT TCTGTAAATA	14400
GATGCATTTT AATTGCAAAA TATCATGCTC AAGATGACAT TGATTTCAAA TTAGATAACA	14460
TTACTATATT AAAAATTAC GTGTGCCTAG GTAGCAAGTT AAAAGGATCT GAAGTTTACT	14520
TAATCCTTAC AATAGGCCCT GCAAATATAC TTCCTGTTTT TGATGTTGTA CAAAATGCTA	14580
AATTGACACT TTCAAGAACT AAAAATTCA TTATGCCTAA AAAAAGTAC AAGGAATCTA	14640
TCGATGCAAA TATTTAAAAGC TTAATACCTT TCCTTTGTTA CCCTATAACA AAAAAGGAA	14700
TTAAGACTTC ATTGTCAAAA TTGAAGAGTG TAGTTAATGG AGATATATTA TCATATTCTA	14760
TAGCTGGACG TAATGAAGTA TTCAGCAACA AGCTTATAAA CCACAAGCAT ATGAATATCC	14820
TAAATGGCT AGATCATGTT TTAAATTTTA GATCAGCTGA ACTTAATTAC AATCATTTAT	14880
ACATGATAGA GTCCACATAT CCTTACTTAA GTGAATTGTT AAATAGTTTA ACAACCAATG	14940
AGCTCAAGAA GCTGATTAAA ATAACAGGTA GTGTGCTATA CAACCTTCCC AACGAACAGT	15000
AGTTTAAAT ATCATTAACA AGTTTGGTCA AATTAGATG CTAACACATC ATTATATTAT	15060
AGTTATTA AAATATACAA ACTTTTCAAT AATTAGCAT ATTGATTCCA AAATTATCAT	15120
TTTAGTCTTA AGGGGTAAA TAAAGTCTA AACTAACA TTATACATGT GCATTCACAA	15180

SUBSTITUTE SHEET (RULE 26)

- 302 -

CACAACGAGA CATTAGTTTT TGACACTTTT TTTCTCGT

15218

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```

Met Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp
1           5           10           15

Ser Tyr Leu Lys Gly Val Ile Ser Phe Ser Glu Cys Asn Ala Leu Gly
20           25           30

Ser Tyr Leu Phe Asn Gly Pro Tyr Leu Lys Asn Asp Tyr Thr Asn Leu
35           40           45

Ile Ser Arg Gln Ser Pro Leu Leu Glu His Met Asn Leu Lys Lys Leu
50           55           60

Thr Ile Thr Gln Ser Leu Ile Ser Arg Tyr His Lys Gly Glu Leu Lys
65           70           75           80

Leu Glu Glu Pro Thr Tyr Phe Gln Ser Leu Leu Met Thr Tyr Lys Ser
85           90           95

Met Ser Ser Ser Glu Gln Ile Ala Thr Thr Asn Leu Leu Lys Lys Ile
100          105          110

Ile Arg Arg Ala Ile Glu Ile Ser Asp Val Lys Val Tyr Ala Ile Leu
115          120          125

Asn Lys Leu Gly Leu Lys Glu Lys Asp Arg Val Lys Pro Asn Asn Asn
130          135          140

Ser Gly Asp Glu Asn Ser Val Leu Thr Thr Ile Ile Lys Asp Asp Ile
145          150          155          160

Leu Ser Ala Val Glu Asn Asn Gln Ser Tyr Thr Asn Ser Asp Lys Ser
165          170          175

His Ser Val Asn Gln Asn Ile Thr Ile Lys Thr Thr Leu Leu Lys Lys

```

SUBSTITUTE SHEET (RULE 26)

- 303 -

180	185	190
Leu Met Cys Ser Met Gln His Pro Pro Ser Trp Leu Ile His Trp Phe		
195	200	205
Asn Leu Tyr Thr Lys Leu Asn Asn Ile Leu Thr Gln Tyr Arg Ser Asn		
210	215	220
Glu Val Lys Ser His Gly Phe Ile Leu Ile Asp Asn Gln Thr Leu Ser		
225	230	235
Gly Phe Gln Phe Ile Leu Asn Gln Tyr Gly Cys Ile Val Tyr His Lys		
245	250	255
Gly Leu Lys Lys Ile Thr Thr Thr Tyr Asn Gln Phe Leu Thr Trp		
260	265	270
Lys Asp Ile Ser Leu Ser Arg Leu Asn Val Cys Leu Ile Thr Trp Ile		
275	280	285
Ser Asn Cys Leu Asn Thr Leu Asn Lys Ser Leu Gly Leu Arg Cys Gly		
290	295	300
Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile		
305	310	315
Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu		
325	330	335
Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe		
340	345	350
Arg Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala		
355	360	365
Ile Lys Ala Gln Lys Asp Leu Leu Ser Arg Val Cys His Thr Leu Leu		
370	375	380
Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu		
385	390	395
Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu		
405	410	415
Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro		
420	425	430
Met Val Asp Glu Arg Gln Ala Met Asp Ser Val Arg Ile Asn Cys Asn		
435	440	445
Glu Thr Lys Phe Tyr Leu Leu Ser Ser Leu Ser Thr Leu Arg Gly Ala		
450	455	460

SUBSTITUTE SHEET (RULE 26)

- 304 -

Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp  
 465 470 475 480  
 Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr  
 485 490 495  
 Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Asn Asp  
 500 505 510  
 Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro  
 515 520 525  
 Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro  
 530 535 540  
 Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser  
 545 550 555 560  
 His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser  
 565 570 575  
 Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe  
 580 585 590  
 Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn  
 595 600 605  
 Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser  
 610 615 620  
 Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln  
 625 630 635 640  
 Ile Leu Ala Glu Lys Met Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro  
 645 650 655  
 Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu  
 660 665 670  
 Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr  
 675 680 685  
 Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe  
 690 695 700  
 Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu  
 705 710 715 720  
 Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr  
 725 730 735

SUBSTITUTE SHEET (RULE 26)



- 305 -

Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe  
 740 745 750  
 Ile Lys Asp His Val Val Asn Leu Asn Glu Val Asp Glu Gln Ser Gly  
 755 760 765  
 Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu  
 770 775 780  
 Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly  
 785 790 795 800  
 Lys Phe Ser Ile Thr Ala Leu Ile Asn Gly Asp Asn Gln Ser Ile Asp  
 805 810 815  
 Ile Ser Lys Pro Val Arg Leu Ile Glu Gly Gln Thr His Ala Gln Ala  
 820 825 830  
 Asp Tyr Leu Leu Ala Leu Asn Ser Leu Lys Leu Leu Tyr Lys Glu Tyr  
 835 840 845  
 Ala Gly Ile Gly His Lys Leu Lys Gly Thr Glu Thr Tyr Ile Ser Arg  
 850 855 860  
 Asp Met Gln Phe Met Ser Lys Thr Ile Gln His Asn Gly Val Tyr Tyr  
 865 870 875 880  
 Pro Ala Ser Ile Lys Lys Val Leu Arg Val Gly Pro Trp Ile Asn Thr  
 885 890 895  
 Ile Leu Asp Asp Phe Lys Val Ser Leu Glu Ser Ile Gly Ser Leu Thr  
 900 905 910  
 Gln Glu Leu Glu Tyr Arg Gly Glu Ser Leu Leu Cys Ser Leu Ile Phe  
 915 920 925  
 Arg Asn Ile Trp Leu Tyr Asn Gln Ile Ala Leu Gln Leu Arg Asn His  
 930 935 940  
 Ala Leu Cys Asn Asn Lys Leu Tyr Leu Asp Ile Leu Lys Val Leu Lys  
 945 950 955 960  
 His Leu Lys Thr Phe Phe Asn Leu Asp Ser Ile Asp Met Ala Leu Ser  
 965 970 975  
 Leu Tyr Met Asn Leu Pro Met Leu Phe Gly Gly Gly Asp Pro Asn Leu  
 980 985 990  
 Leu Tyr Arg Ser Phe Tyr Arg Arg Thr Pro Asp Phe Leu Thr Glu Ala  
 995 1000 1005  
 Ile Val His Ser Val Phe Val Leu Ser Tyr Tyr Thr Gly His Asp Leu

SUBSTITUTE SHEET (RULE 26)

- 306 -

1010	1015	1020
Gln Asp Lys Leu Gln Asp Leu Pro Asp Asp Arg Leu Asn Lys Phe Leu 1025	1030	1035 1040
Thr Cys Val Ile Thr Phe Asp Lys Asn Pro Asn Ala Glu Phe Val Thr 1045	1050	1055
Leu Met Arg Asp Pro Gln Ala Leu Gly Ser Glu Arg Gln Ala Lys Ile 1060	1065	1070
Thr Ser Glu Ile Asn Arg Leu Ala Val Thr Glu Val Leu Ser Ile Ala 1075	1080	1085
Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu 1090	1095	1100
Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His 1105	1110	1115 1120
Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys 1125	1130	1135
Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu 1140	1145	1150
Lys Thr Ser Ala Ile Asp Thr Thr Asp Ile Asn Arg Ala Thr Asp Met 1155	1160	1165
Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys 1170	1175	1180
Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr 1185	1190	1195 1200
Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile 1205	1210	1215
Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asp Ile Lys Tyr 1220	1225	1230
Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val 1235	1240	1245
Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly 1250	1255	1260
Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val 1265	1270	1275 1280
Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp 1285	1290	1295

SUBSTITUTE SHEET (RULE 26)

- 307 -

Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu  
 1300 1305 1310

Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe  
 1315 1320 1325

Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser  
 1330 1335 1340

Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn  
 1345 1350 1355 1360

Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr  
 1365 1370 1375

Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly  
 1380 1385 1390

Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn  
 1395 1400 1405

Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro  
 1410 1415 1420

Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile  
 1425 1430 1435 1440

Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr  
 1445 1450 1455

Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile  
 1460 1465 1470

Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn  
 1475 1480 1485

Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile  
 1490 1495 1500

Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu  
 1505 1510 1515 1520

Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn  
 1525 1530 1535

Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala  
 1540 1545 1550

Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu  
 1555 1560 1565

SUBSTITUTE SHEET (RULE 26)

- 308 -

Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu  
 1570 1575 1580

Gln Lys Val Ile Lys Tyr Ile Val Asn Gln Asp Thr Ser Leu Arg Arg  
 1585 1590 1595 1600

Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asn  
 1605 1610 1615

Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His  
 1620 1625 1630

Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met  
 1635 1640 1645

Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe  
 1650 1655 1660

Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe  
 1665 1670 1675 1680

Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser  
 1685 1690 1695

Glu Leu Glu Asp Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr  
 1700 1705 1710

Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys  
 1715 1720 1725

Pro Lys Phe Cys Ile Ser Gly Asn Thr Glu Ser Met Met Met Ser Thr  
 1730 1735 1740

Phe Ser Ser Lys Met His Ile Lys Ser Ser Thr Val Thr Thr Arg Phe  
 1745 1750 1755 1760

Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile  
 1765 1770 1775

Asp Lys Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu  
 1780 1785 1790

Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser  
 1795 1800 1805

Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val  
 1810 1815 1820

Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp  
 1825 1830 1835 1840

Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala

SUBSTITUTE SHEET (RULE 26)

- 309 -

1845	1850	1855
Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg 1860	1865	1870
Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile 1875	1880	1885
Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu 1890	1895	1900
Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser 1905	1910	1915 1920
Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp 1925	1930	1935
Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp 1940	1945	1950
Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys 1955	1960	1965
Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu 1970	1975	1980
Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu 1985	1990	1995 2000
Lys Gly Ser Glu Val Tyr Leu Ile Leu Thr Ile Gly Pro Ala Asn Ile 2005	2010	2015
Leu Pro Val Phe Asp Val Val Gln Asn Ala Lys Leu Thr Leu Ser Arg 2020	2025	2030
Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp 2035	2040	2045
Ala Asn Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys 2050	2055	2060
Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Asn Gly 2065	2070	2075 2080
Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn 2085	2090	2095
Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His 2100	2105	2110
Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met 2115	2120	2125

SUBSTITUTE SHEET (RULE 26)

- 310 -

Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr  
 2130 2135 2140

Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr  
 2145 2150 2155 2160

Asn Leu Pro Asn Glu Gln  
 2165

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15229 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ACGCGAAAAA ATGCGTACTA CAACTTGCA CATTGCGAAA AAATGGGGCA AATAAGAATT	60
TGATAAGTGC TATTTAAATC TAACCTTTTC AATCAGAAAT GGGGTGCAAT TCACTGAGCA	120
TGATAAAGGT TAGATTACAA AATTTATTTG ACAATGACGA AGTAGCATTG TTAAAAATAA	180
CATGTTATAC TGACAAATTA ATTCTTCTGA CCAATGCATT AGCCAAAGCA GTAATACATA	240
CAATTAAATT AAACGGCATA GTTTTATAC ATGTTATAAC AAGCAGTGAA GTGTGCCCTG	300
ACAACAATAT TGTAGTGAAA TCTAACTTTA CAACAATGCC AATATTACAA AACGGAGGAT	360
ACATATGGGA ATTGATTGAG TTGACACACT GCTCTCAATC AAATGGTCTA ATGGATGATA	420
ATTGTGAAAT CAAATTTTCT AAAAGACTAA GTGACTCAGT AATGACTAAT TATATGAATC	480
AAATATCTGA TTTACTTGGG CTTGATCTCA ATTCATGAAT TATGTTTAGT CTAATTTAAT	540
AGACATGTGT TTATCACCAT TTTAGTTAAT ATAAACCTC ATCAAAGGGA AATGGGGCAA	600
ATAAACTCAC CTAATCAGTC AAACCATGAG CACTACAAAT GACAACACTA CTATGCAAAG	660
ATTGATGATC ACAGACATGA GACCCCTGTC GATGGAATCA ATAATAACAT CTCTCACCAA	720
AGAAATCATA ACACACAAAT TCATATACTT GATAACAAT GAATGTATTG TAAGAAAAC	780
TGATGAAAGA CAAGCTACAT TTACATTCTT AGTCAATTAT GAGATGAAGC TATTGCACAA	840

- 311 -

AGTAGGGAGT ACCAAATACA AGAAATACAC TGAATATAAT ACAAATATG GCACTTTCCC 900  
CATGCCTATA TTTATCAATC ATGACGGGTT TCTAGAATGT ATTGGCATT A GCCTACAAA 960  
ACACACTCCT ATAATATACA AATATGACCT CAACCCGTAA ATTCCAACAA AAAACTAACC 1020  
CATCCAAACT AAGCTATTCC TCAAACAACA GTGCTCAACA GTTAAGAAGG AGCTAATCCA 1080  
TTT TAGTAAT TAAAAATAAA GGCAGAGCCA ATAACATAAA TTGGGGCAAA TACAAAGATG 1140  
GCTCTTAGCA AAGTCAAGTT AAATGATACA TTAAATAAGG ATCAGCTGCT GTCATCCAGC 1200  
AAATACACTA TTCAACGTAG TACAGGAGAT AATATTGACA CTCCCAATTA TGATGTGCAA 1260  
AAACACCTAA ACAAACATG TGGTATGCTA TTAATCACTG AAGATGCAAA TCATAAATTC 1320  
ACAGGATTAA TAGGTATGTT ATATGCTATG TCCAGGTTAG GAAGGGAAGA CACTATAAAG 1380  
ATACTTAAAG ATGCTGGATA TCATGTTAAA GCTAATGGAG TAGATATAAC AACATATCGT 1440  
CAAGATATAA ACGGAAAGGA AATGAAATTC GAAGTATTAA CATTATCAAG CTTGACATCA 1500  
GAAATACAAG TCAATATTGA GATAGAATCT AGAAAGTCCT ACAAAAAAAT GCTAAAAGAG 1560  
ATGGGAGAAG TGGCTCCAGA ATATAGGCAT GATTCTCCAG ACTGTGGGAT GATAATACTG 1620  
TGTATAGCTG CACTTGTAAT AACCAAGTTA GCAGCAGGAG ATAGATCAGG TCTTACAGCA 1680  
GTAATTAGGA GGGCAAACAA TGTCTTAAA AACGAAATAA AACGCTACAA GGGCCTCATA 1740  
CCAAAGGATA TAGCTAACAG TTTTATGAA GTGTTTGAAA AACACCCTCA TCTTATAGAT 1800  
GTTTTTGTGC ACTTTGGCAT TGCACAATCA TCCACAAGAG GGGGTAGTAG AGTTGAAGGA 1860  
ATCTTTGCAG GATTATTTAT GAATGCCTAT GGTTCAGGGC AAGTAATGCT AAGATGGGGA 1920  
GTTCTAGCCA AATCTGTAAA AAATATCATG CTAGGACATG CTAGTGTCCA GGCAGAAATG 1980  
GAACAAGTTG TGGAAGTTA TGAGTATGCA CAGAAGTTGG GAGGAGAAGC TGGATTCTAC 2040  
CATATATTGA ACAATCCAAA AGCATCATTG CTGTCATTAA CTCAATTTCC TAACTTCTCA 2100  
AGTGTGGTCC TAGGCAATGC AGCAGGTCTA GGCATAATGG GAGAGTATAG AGGTACACCA 2160  
AGAAACCAAG ATCTATATGA TGCAGCCAAA GCATATGCAG AGCAACTCAA AGAAAATGGA 2220  
GTAATAAACT ACAGTGTATT AGACTTAACA GCAGAAGAAT TGGAAGCCAT AAAGCATCAA 2280  
CTCAACCCCA AAGAAGATGA TG TAGAGCTT TAAGTTAACA AAAAAACGG GGCAAAATAAG 2340  
TCAACATGGA GAAGTTTGCA CCTGAATTTT ATGGAGAAGA TGCAAACAAC AAAGCTACCA 2400

SUBSTITUTE SHEET (RULE 26)

- 312 -

AATTCCTAGA ATCAATAAAG GGCAAGTTTG CATCATCCAA AGATCCTAAG AAGAAAGATA 2460  
GCATAATATC TGTAACTCA ATAGATATAG AAGTAACTAA AGAGAGCCCG ATAACATCTG 2520  
GCACCAACAT CATCAATCCA ATAAGTGAAG CTGATAGTAC CCCAGAAGCT AAAGCCAACT 2580  
ACCCAAGAAA ACCCCTAGTA AGCTTCAAAG AAGATCTCAC CCCAAGTGAC AACCCCTTTT 2640  
CTAAGTTGTA CAAAGAAACA ATAGAAACAT TTGATAACAA TGAAGAAGAA TCTAGCTACT 2700  
CATATGAAGA AATAAATGAT CAAACAAATG ACAACATTAC AGCAAGACTA GATAGAATTG 2760  
ATGAAAAATT AAGTGAAATA TTAGGAATGC TCCATACATT AGTAGTTGCA AGTGCAGGAC 2820  
CCACCTCAGC TCGCGATGGA ATAAGAGATG CTATGGTTGG TCTAAGAGAA GAAATGATAG 2880  
AAAAATAAG AGCGGAAGCA TTAATGACCA ATGATAGGTT AGAGGCTATG GCAAGACTTA 2940  
GGAATGAGGA AAGCGAAAAA ATGGCAAAAG ACACCTCAGA TGAAGTGTCT CTTAATCCAA 3000  
CTTCAAAAAA ATTGAGTAAT TTGTTGGAAG ACAACGATAG TGACAATGAT CTATCACTTG 3060  
ATGATTTTTG ATCAGTGATC AACTCACTCA GCAATCAACA ACATCAATGA AACAGACATC 3120  
AATCCATTGA ATCAACTGCC AGACTGAACA CACAAACGTC CATCAGCAGA ACTACCAACC 3180  
AATCAATCAA CCAATTGATC AATCAGCGAC CTAACAAAAT TAACAATATA GTAACAAAAA 3240  
AAGAACAAGA TGGGGCAAAT ATGGAAACAT ACGTGAACAA GCTTCACGAG GGCTCCACAT 3300  
ACACAGCAGC TGTTCACTAC AATGTTCTAG AAAAAGATGA TGATCCTGCA TCACTAACAA 3360  
TATGGGTGCC TATGTTCCAG TCATCTGTGC CAGCAGACTT GCTCATAAAA GAACTTGCAA 3420  
GCATCAACAT ACTAGTGAAG CAGATCTCCA CGCCCAAAGG ACCTTCACTA CGAGTCACGA 3480  
TTAACTCAAG AAGTGCTGTG CTGGCACAAA TGCCTAGTAG TTTTATCATA AGTGCAAATG 3540  
TATCATTAGA TGAAAGAAGC AAATTAGCAT ATGATGTAAC TACACCTGTG GAAATCAAAG 3600  
CATGCAGTCT AACATGCTTA AAAGTAAAAA GTATGTTAAC TACAGTCAAA GATCTTACCA 3660  
TGAAACATT CAATCCCACT CATGAGATTA TTGCTCTATG TGAATTTGAA AATATTATGA 3720  
CATCAAAAAG AGTAATAATA CCAACCTATC TAAGATCAAT TAGTGTCAAA AACAAAGGACC 3780  
TGAACTCACT AGAAAATATA GCAACCACCG AATTCAAAAA TGCTATCACC AATGCGAAAA 3840  
TTATTCCCTA TGCAGGATTA GTATTAGTTA TCACAGTTAC TGACAATAAA GGAGCATTCA 3900  
AATATATCAA GCCACAGAGT CAATTTATAG TAGATCTTGG GGCCTACCTA GAAAAAGAGA 3960

SUBSTITUTE SHEET (RULE 26)



- 313 -

GCATATATTA TGTGACTACA AATTGGAAGC ATACAGCTAC ACGTTTTTCA ATCAAACCAC 4020  
TAGAGGATTA AACTTAATTA TCAACACTAA ATGACAGGTC CACATATATC TTCAAACTAT 4080  
ACATTATATC CAAACATCAT GAGCATTAC ACTACACACT TTTACCATAT AAATCAATCT 4140  
CATTTAAAAT CCAAAATTAC TTCCAGCTAT CATCTGTTAG ACCTAGAGTG CGAATAGGTA 4200  
AATAAAACCA AAATATGGGG TAAATAGACA TTAGTTAGAG TTCAATCAAT CTCAACAACC 4260  
ATTTATACCG CCAATTCAGT ACATATACTA TAAATCTCAA AATGGGAAAT ACATCCATCA 4320  
CAATAGAATT CACAAGCAAA TTTTGGCCTT ATTTTACACT AATACATATG ATCTTAACTC 4380  
TAATCTCTTT ACTAATTATA ATCACTATTA TGATTGCAAT ACTAAATAAG CTAAGTGAAC 4440  
ATAAAACATT CTGCAACAAA ACTCTTGAAC TAGGACAGAT GTATCAAATC AACACATAGT 4500  
GTTCTACCAT TATGCTGTGT CAAATTATAA TCTTGTATAT ATAAACAAAC AAATCCAATC 4560  
TTCTCACAGA GTCATGGTGG CGCAAAACCA CGCCAACCAT CATGATAGCA TAGAGTAGTT 4620  
ATTTAAAAAT TAACATAATG ATGAATTATT GGTATGAGAT CAGGAACAAC ATTGGGGCAA 4680  
ATGCAGCCAT GTCCAAGCAC AAGAATCGGC GCACTGCCGG GACTCTAGAA AGGACCTGGG 4740  
ATACTCTTAA TCATCTAATT GTAATATCCT CTTGTTTATA CAGATTAAAT TTAAAATCTA 4800  
TAGCACAAAT AGCACTGTCA GTTTTGGCAA TGATAATCTC AACCTCTCTC ATAATTGCAG 4860  
CCATAATATT CATCATCTCT GCCAATCACA AAGTTACACT AACAACGGTT ACAGTTCAAA 4920  
CAATAAAAAA CCACACTGAA AAAACATCT CCACCTACCT TACTCAAGTC CCACCAGAAA 4980  
GGGTCAACTC ATCCAAACAA CCCACAACCA CATCACCAT CCACACAAAT TCAGCCACAA 5040  
TATCACCAA TACAAAATCA GAAACACACC ATACAACAGC ACAAACCAA GGCAGAATCA 5100  
CCACTTCAAC ACAGACCAAC AAGCCAAGCA CAAAATCAG TTCAAAAAAT CCACCAAAAA 5160  
AACCAAAAGA TGATTACCAT TTTGAAGTGT TCAATTTTGT TCCCTGTAGT ATATGTGGTA 5220  
ATAATCAACT CTGCAAATCC ATCTGCAAAA CAATACCAAG CAACAAACCA AAGAAAAAAC 5280  
CAACCATCAA ACCCACAAC AAACCAACCA CCAAAACCAC AAACAAAAGA GACCCCAAAA 5340  
CACCAGCCAA AATGCCAAA AAAGAAATCA TCACCAACCC AGCAAAAAAA CCAACCCTCA 5400  
AGACCACAGA AAGAGACACC AGCATTTCAC AATCCACCGT GCTCGACACA ATCACTCCAA 5460  
AATACACAAT CCAACAGCAA TCCCTCCACT CAACCACCTC CGAAAACACA CCCAGCTCCA 5520

SUBSTITUTE SHEET (RULE 26)

- 314 -

CACAAATACC CACAGCATCC GAGCCCTCCA CATTAAATCC TAATTA AAAA ACCTAGTCAC	5580
ATGCTTAGTT ATTCAAAAAC TACATCTTAG CAGAGAACCG TGATCTATCA AGCAAGAACA	5640
AAATTAAACC TGGGGCAAAT AACCATGGAG TTGCTGATCC ACAGGTCAAG TGCAATCTTC	5700
CTAACTCTTG CTGTTAATGC ATTGTACCTC ACCTCAAGTC AGAACATAAC TGAGGAGTTT	5760
TACCAATCGA CATGTAGTGC AGTTAGCAGA GGTATTTTA GTGCTTTAAG AACAGGTTGG	5820
TATACCAGTG TCATAACAAT AGAATTAAGT AATATAAAAG AAACCAAATG CAATGGAACT	5880
GACACTAAAG TAAAACTTAT AAAACAAGAA TTAGATAAGT ATAAGAATGC AGTAACAGAA	5940
TTACAGCTAC TTATGCAAAA CACGCCAGCT GCCAACAACC GGGCCAGAAG AGAAGCACCA	6000
CAGTACATGA ACTACACAAT CAATACCACA AAAAACCTAA ATGTATCAAT AAGCAAGAAA	6060
AGGAAACGAA GATTTCTGGG CTTCTTGTTA GGTGTAGGAT CTGCAATAGC AAGTGGTATA	6120
GCTGTATCCA AAGTTTTACA CCTTGAAGGA GAAGTGAACA AAATCAAAA TGCTTTGTTG	6180
TCTACAAACA AAGCTGTAGT CAGTCTATCA AATGGGGTCA GTGTTTTAAC CAGCAAAGTG	6240
TTAGATCTCA AGAATTACAT AAATAACCGA ATATTACCCA TAGTAAATCA ACAGAGCTGT	6300
CGCATCTCCA ACATTGAAAC AGTTATAGAA TTCCAGCAGA AGAATAGCAG ATTGTTGGAA	6360
ATCACCAGAG AATTTAGTGT TAATGCAGGT GTAACAACAC CTTTAAGCAC TTACATGTTA	6420
ACAAACAGTG AGTTACTATC ATTGATCAAT GATATGCCTA TAACAAATGA CCAGAAAAAA	6480
TTAATGTCAA GCAATGTTCA GATAGTAAGG CAACAAAGTT ATTCTATCAT GTCTATAATA	6540
AAGGAAGAAG TCCTTGCAAT TGTGTACAG CTACCTATCT ATGGTGTAAT AGATACACCT	6600
TGCTGGAAAT TACACACATC ACCTCTATGC ACCACCAACA TCAAAGAAGG ATCAAATATT	6660
TGTTTAACAA GGACTGATAG AGGATGGTAT TGTGATAATG CAGGATCAGT ATCCTTCTTC	6720
CCACAGGCTG ATACTTGCAA AGTACAGTCC AATCGAGTAT TTTGTGACAC TATGAACAGT	6780
TTACATTAC CAAGTGAAGT CAGCCTTTGT AACACTGACA TATTCAATTC CAAGTATGAC	6840
TGCAAAATTA TGACATCAAA AACAGACATA AGCAGCTCAG TAATTACTTC TCTTGAGCT	6900
ATAGTGTCAT GCTATGGAAA AACTAAATGC ACTGCATCCA ATAAAAATCG TGGGATTATA	6960
AAGACATTTT CTAATGGTTG TGAATATGTG TCAAACAAAG GAGTAGATAC TGTGTCAGTG	7020
GGCAACACTT TATACTATGT AAACAAGCTG GAAGGCAAAA ACCTTTATGT AAAAGGGGAA	7080

SUBSTITUTE SHEET (RULE 26)

- 315 -

CCTATAATAA ATTACTATGA TCCTCTAGTG TTTCCTTCTG ATGAGTTTGA TGCATCAATA	7140
TCTCAAGTCA ATGAAAAAAT CAATCAAAGT TTAGCTTTTA TTCGTAGATC TGATGAATTA	7200
CTACATAATG TAAATACTGG CAAATCTACT ACAAATATTA TGATAACTAC AATTATTATA	7260
GTAATCATTG TAGTATTGTT ATCATTAAATA GCTATTGGTT TACTGTTGTA TTGCAAAGCC	7320
AAAAACACAC CAGTTTACT AAGCAAAGAC CAACTAAGTG GAATCAATAA TATTGCATTG	7380
AGCAAATAGA CAAAAAATA CTTAATCATG TTTCAACAAC AATCTGCTGA CCACCAATCC	7440
CAAATCAACT TAACAACAAA TATTTCAACA TCATAGCACA GGCTGAATCA TTTCCTCATA	7500
TCATGCTACC TACACAACTA AGCTAGATCT TCAACTCATA GTTACATAAA AACCCCAAGT	7560
ATCACAATCA AACACTAAAT CGACACATCA TTCACAAAAT TAACAAGTGG GGCAAATATG	7620
TCGCGAAGAA ATCCTTGTA AATTGAGATT AGAGGTCATT GCTTGAATGG TAGAAGATGT	7680
CACTACAGTC ATAATTATTT TGAATGGCCT CCTCATGCAT TACTAGTGAG GCAAACTTC	7740
ATGTTAAACA AGATACTTAA GTCAATGGAC AAAAGCATAG AACTTTGTC GGAAATAAGT	7800
GGAGCTGCTG AACTGGATAG AACAGAAGAA TATGCTCTTG GTATAGTTGG AGTGCTAGAG	7860
AGTTACATAG GATCAATAAA CAACATAACA AAACAATCAG CATGTGTTGC TATGAGTAAA	7920
CTTCTTATTG AGATCAACAG TGATGACATT AAAAACTGA GAGATAACGA AGAACCCAAT	7980
TCGCCTAAGA TAAGAGTGTA CAATACTGTT ATATCATACA TTGAGAGCAA TAGAAAAAAC	8040
AACAAGCAAA CCATCCATCT GCTCAAAGA CTACCAGCAG ACGTGCTGAA GAAGACAATA	8100
AAGAACACAT TAGATATCCA CAAAAGCATA ACCATAAGCA ACTCAAAGA GTCAACCGTG	8160
AATGATCAAA ATGACCAAAC CAAAAATAAT GATATTACCG GATAAATATC CTTGTAGTAT	8220
ATCATCCATA TTGATTTCAA GTGAAAGCAT GATTGCTACA TTCAATCATA AAAACATATT	8280
ACAATTTAAC CATAACCATT TGGATAACCA CCAGTGTTTA TTAAATCATA TATTTGATGA	8340
AATTCATTGG ACACCTAAAA ACTTATTAGA TGCCACTCAA CAATTTCTCC AACATCTTAA	8400
CATCCCTGAA GATATATATA CAGTATATAT ATTAGTGTCA TAATGCTTGA CCATAACAAT	8460
TTTATATCAT TCAACCATAA AACAACCTTA ATAAGGTTAT GGGACAAAAT GGATCCCATT	8520
ATTAATGGAA ACTCTGCCAA TGTGTATCTA ACTGATAGTT ATCTAAAAGG TGTTATCTCT	8580
TTTTCAGAAT GTAATGCTTT AGGGAGTTAC CTTTTTAACG GCCCCTATCT TAAAAATGAT	8640

SUBSTITUTE SHEET (RULE 26)

- 316 -

TACACCAACT TAATTAGTAG ACAAAGCCCA CTACTAGAGC ATATGAATCT AAAAAAACTA	8700
ACTATAACAC AGTCATTAAT ATCTAGATAT CATAAAGGTG AACTGAAGTT AGAAGAACCA	8760
ACTTATTTCC AGTCATTACT TATGACATAT AAAAGTATGT CCTCGTCTGA ACAAATTGCT	8820
ACAACTAATT TACTTAAAAA AATAATACGA AGAGCTATAG AAATAAGTGA TGTAAGGTG	8880
TACGCCATCT TGAATAAACT GGGACTAAAG GAAAAGGACA GAGTTAAGCC CAACAATAAT	8940
TCAGGTGATG AAAACTCAGT TCTTACAACC ATAATCAAAG ATGATATACT TTCAGCTGTG	9000
GAAAACAATC AATCATATAC AAATTCAGAC AAAAATCATT CAGTAAATCA AAATATCACT	9060
ATCAAAACAA CACTCTTGAA AAAATTGATG TGTTCATGC AACATCCTCC ATCATGGTTA	9120
ATACACTGGT TCAATTTATA TACAAAATTA AATAACATAT TAACACAATA TCGATCAAAT	9180
GAGGTAAAAA GTCATGGGTT TATATTAATA GATAATCAAA CTTTAAGTGA TTTTCAGTTT	9240
ATTTTAAATC AATATGGTTG TATCGTTTAT CATAAAGGAC TCAAAAAAAT CACAACACT	9300
ACTTACAATC AATTTTGGAC ATGGAAAGAC ATCAGCCTTA GCAGATTAAA TGTTTGCTTA	9360
ATTACTTGGA TAAGTAATTG TTTAAATACA TTAATAAAAA GCTTAGGGCT GAGATGTGGA	9420
TTCAATAATG TTGTGTTATC ACAACTATTT CTTTATGGAG ATTGTATACT GAAATTATTC	9480
CATAATGAAG GCTTCTACAT AATAAAAGAA GTAGAGGGAT TTATTATGTC TTTAATTCTA	9540
AACATAACAG AAGAAGATCA ATTTAGGAAA CGATTTTATA ATAGCATGCT AAATAACATC	9600
ACAGATGCAG CTATTAAGGC TCAAAAAAAC CTACTATCAA GAGTATGTCA CACTTTATTA	9660
GACAAGACAG TGTCTGATAA TATCATAAAT GGTAAATGGA TAATCCTATT AAGTAAATTT	9720
CTTAAATTGA TTAAGCTTGC AGGTGATAAT AATCTCAATA ACTTGAGTGA GCTTTATTTT	9780
CTCTTCAGAA TCTTTGGACA TCCAATGGTC GATGAAAGAC AAGCAATGGA TGCTGTAAGA	9840
ATTAAGTGTA ATGAAACCAA GTTCTACTTA TTAAGTAATC TAAGTACGTT AAGAGGTGCT	9900
TTCATTTATA GAATCATAAA GGGGTTTGTA AATACCTACA ACAGATGGCC CACTTTAAGG	9960
AATGCTATTG TTCTACCTCT AAGATGGTTG AACTATTATA AACTTAATAC TTATCCATCT	10020
CTACTTGAAA TCACAGAGAA AGATTTGATT ATTTTATCAG GATTGCGGTT CTATCGTGAG	10080
TTTCATCTGC CTAAAAAGT GGATCTTGAA ATGATAATAA ATGACAAAGC CATTTACCTT	10140
CCAAAAGATT TAATATGGAC TAGTTTTTCCT AGAAATTACA TGCCATCACA TATACAAAAT	10200

SUBSTITUTE SHEET (RULE 26)

- 317 -

TATATAGAAC ATGAAAAGTT GAAGTTCTCT GAAAGTGACA GATCAAGAAG AGTACTAGAG 10260  
TATTACTTGA GAGATAATAA ATTCAATGAA TGCGATCTAT ACAATTGTGT GGTCAATCAA 10320  
AGCTATCTCA ACAACTCTAA CCATGTGGTA TCACTAACTG GTAAAGAAAG AGAGCTCAGT 10380  
GTAGGTAGAA TGTTTGCTAT GCAACCAGGT ATGTTTAGGC AAATTCAAAT CTTAGCAGAG 10440  
AAAATGATAG CCGAAAATAT TTTACAATTC TTCCCTGAGA GTTTGACAAG ATATGGTGAT 10500  
CTAGAGCTTC AAAAGATATT AGAATTAAAA GCAGGAATAA GCAACAAGTC AAATCGTTAT 10560  
AATGATAACT ACAACAATTA TATCAGTAAA TGTTCTATCA TTACAGACCT TAGCAAATTC 10620  
AATCAAGCAT TTAGATATGA AACATCATGT ATCTGCAGTG ATGTATTAGA TGAAGTGCAT 10680  
GGAGTACAAT CTCTGTTCTC TTGGTTGCAT TTAACAATAC CTCTTGTCAC AATAATATGT 10740  
ACATATAGAC ATGCACCTCC TTTTATAAAG GATCATGTTG TTAATCTTAA TAAAGTTGAT 10800  
GAACAAAGTG GATTATACAG ATATCATATG GGTGGTATTG AAGGCTGGTG TCAAAAAC TG 10860  
TGGACCATTG AAGCTATATC ATTATTAGAT CTAATATCTC TCAAAGGGAA ATTCTCTATC 10920  
ACAGCTCTAA TAAATGGTGA TAATCAGTCA ATTGATATAA GTAAACCAAGT TAGACTTATA 10980  
GAGGGTCAGA CCCATGCTCA AGCAGATTAT TTGTTAGCAT TAAATAGCCT TAAATTGCTA 11040  
TATAAAGAGT ATGCGGGCAT AGGCCACAAG CTCAAGGGAA CAGAGACCTA TATATCCCGA 11100  
GATATGCAAT TCATGAGCAA AACAATCCAG CACAATGGAG TGTACTATCC AGCCAGTATC 11160  
AAAAAAGTCC TGAGAGTAGG TCCATGGATA AATACAATAC TTGATGATT TAAAGTTAGT 11220  
TTAGAATCTA TAGGTAGCTT AACACAGGAG TTAGAATATA GAGGAGAGAG CTTATTATGC 11280  
AGTTTAATAT TTAGGAACAT TTGGTTATAC AATCAAATTG CTTTGCAACT CCGAAATCAT 11340  
GCATTATGTC ACAATAAGCT ATATTTAGAT ATATTGAAAG TATTAACA CTTAAAAACT 11400  
TTTTTTAATC TTGATAGTAT TGATATGGCT TTAACATTGT ATATGAATT GCCTATGCTG 11460  
TTTGGTGGTG GTGATCCTAA TTTGTTATAT CGAAGCTTTT ATAGGAGAAC TCCAGACTTC 11520  
CTTACAGAAG CTATAGTACA TTCAGTGTTC GTGTTGAGCT ATTATACTGG TCACGATTTA 11580  
CAAGATAAGC TCCAGGATCT TCCAGATGAT AGACTGAACA AATTCTTGAC ATGTATCATC 11640  
ACGTTTGATA AAAATCCCAA TGCCGAGTTT GTAACATTGA TGAGAGATCC ACAGGCTTTA 11700  
GGGTCTGAAA GGCAAGCAAA AATTACTAGT GAGATTAAATA GATTAGCAGT GACAGAAGTC 11760

SUBSTITUTE SHEET (RULE 26)

- 318 -

TTAAGTATAG CTCCAAACAA AATATTTTCT AAAAGTGCAC AACATTATAC TACCACTGAG	11820
ATTGATCTAA ATGATATTAT GCAAAATATA GAACCAACTT ACCCTCATGG ATTAAGAGTT	11880
GTTTATGAAA GTTTACCTTT TTATAAAGCA GAAAAAATAG TTAATCTTAT ATCAGGAACA	11940
AAATCCATAA CTAATATACT TGAAAAACA TCAGCAATAG ATTCAACTGA TATTAATAGG	12000
GCTACTGATA TGATGAGGAA AAATATAACT TTACTTATAA GGATACTTCC ACTAGATTGT	12060
AACAAAGACA AAAGAGAGTT ATTAAGTTTA GAAATCTTA GTATAACTGA ATTAAGCAAG	12120
TATGTAAGAG AAAGATCTTG GTCGTTATCC AATATAGTAG GAGTAACATC GCCAAGTATT	12180
ATGTTCAACA TGGACATTAA ATATACAACT AGCACTATAG CCAGTGGTAT AATTATAGAA	12240
AAATATAATG TTAATAGTTT AACTCGTGGT GAAAGAGGAC CTACTAAGCC ATGGGTAGGT	12300
TCATCTACGC AGGAGAAAAA AACAAATGCCA GTGTACAATA GACAAGTTTT AACC AAAAAG	12360
CAAAGAGACC AAATAGATTT ATTAGCAAAA TTAGACTGGG TATATGCATC CATAGACAAC	12420
AAAGATGAAT TCATGGAAGA ACTGAGTACT GGAACACTTG GACTGTCATA TGAGAAAGCC	12480
AAAAAATTGT TTCCACAATA TCTAAGTGTC AATTATTTAC ACCGCTTAAC AGTCAGTAGT	12540
AGACCATGTG AATTCCCTGC ATCAATACCA GCTTATAGAA CAACAAATTA TCATTTTCGAT	12600
ACTAGTCCTA TCAACCATGT ATTAACAGAA AAGTATGGAG ATGAAGATAT CGACATTGTG	12660
TTTCAAAATT GCATAAGTTT TGGTCTTAGC TTAATGTCGG TTGTGGAACA ATTCACAAAC	12720
ATATGTCCTA ATAGAATTAT TCTCATACCG AAGCTGAATG AGATACATTT GATGAAACCT	12780
CCTATATTTA CAGGAGATGT TGATATCATC AAGTTGAAGC AAGTGATACA AAAACAGCAC	12840
ATGTTCCCTAC CAGATAAAAT AAGTTTAACC CAATATGTAG AATTATTCCT AAGTAACAAA	12900
GCACTTAAAT CTGGATCTCA CATCAACTCT AATTTAATAT TAGTACATAA AATGTCTGAT	12960
TATTTTCATA ATGCTTATAT TTTAAGTACT AATTTAGCTG GACATTGGAT TCTGATTATT	13020
CAACTTATGA AGGATTCAAA AGGTATTTTT GAAAAAGATT GGGGAGAGGG GTATATAACT	13080
GATCATATGT TCATTAATTT GAATGTTTTT TTTAATGCTT ATAAGACTTA TTTGCTATGT	13140
TTTCATAAAG GTTATGGTAA AGCAAAATTA GAATGTGATA TGAACACTTC AGATCTTCTT	13200
TGTGTTTTGG AGCTAATAGA CAGTAGCTAC TGGAAATCTA TGTCTAAAGT TTTCTAGAA	13260
CAAAAAGTCA TAAAATACAT AATCAATCAA GACACAAGTT TGCATAGAAT AAAAGGTTGT	13320

SUBSTITUTE SHEET (RULE 26)

- 319 -

CATAGTTTTA AGTTATGGTT TTAAAAACGC CTTAATAATG CTAAATTTAC CGTATGCCCT 13380  
TGGGTTGTTA ACATAGATTA TCACCCAACA CACATGAAAG CTATATTATC TTACATAGAT 13440  
TTAGTTAGAA TGGGGTTAAT AAATGTAGAT AAATTAACCA TAAAAAATAA AAATAAATTC 13500  
AATGATGAAT TTTACACATC AAATCTCTTT TACATTAGTT ATAACTTTTC AGATAACACT 13560  
CATTTGCTAA CAAAACAAAT AAGAATTGCT AATTCAGAAT TAGAAAATAA TTATAACAAA 13620  
CTATATCACC CAACCCCAAGA AACTTTAGAA AATATGTCAT TAATTCCTGT CAAAAGTAAT 13680  
AATAGTAATA AACCTAAAT TGGTATAAGT GGAAATACCG AATCTATGAT GACGTCAACA 13740  
TTCTCCAATA AAACGCATAT TAAATCTTCC GCTGTTATTA CAAGATTCAA TTATAGTAAA 13800  
CAAGACTTGT ACAATTTATT TCCAATTGTC GTGATAGACA GGATTATAGA TCATTCAGGT 13860  
AATACAGCAA AATCTAACCA ACTCTACACT ACCACTTCAC ATCAGACATC TTTAGTAAGG 13920  
AATAGTGCAT CACTTTATTG CATGCTTCCT TGGCATCATG TCAATAGATT TAACTTTGTA 13980  
TTTAGTTCCA CAGGATGCAA GATCAGTATA GAGTATATTT TAAAAGATCT TAAGATTAAA 14040  
GACCCAGTT GTATAGCATT CATAGGTGAA GGAGCTGGTA ACTTATTATT ACGTACAGTA 14100  
GTAGAACTTC ATCCAGACAT AAGATACATT TACAGAAGTT TAAAAGATTG CAATGATCAT 14160  
AGTTTACCTA TTGAATTCT AAGGTTATAC AACGGGCATA TAAACATAGA TTATGGTGAG 14220  
AATTTAACCA TTCCTGCTAC AGATGCAACT AATAACATTC ATTGGTCTTA TTTACATATA 14280  
AAATTTGCAG AACCTATTAG CATTTTGTG TCGGATGCTG AATTACCTGT TACAGCCAAT 14340  
TGGAGTAAAA TTATAATTGA ATGGAGTAAG CATGTAAGAA AGTGCAAGTA CTGTTCTCT 14400  
GTAAATAGAT GCATTTTAAT TGCAAAATAT CATGCCCAAG ATGATATTGA TTTCAAATTA 14460  
GATAACATTA CTATATTAAG AACCTACGTG TGCCTAGGTA GCAAGTTAAA AGGATCTGAA 14520  
GTTTACTTAG TCCTTACAAT AGGCCCTGCA AATATACTTC CTGTTTTTAA TGTGTGCAA 14580  
AATGCTAAAT TGATTCTTTC AAGGACTAAA AATTCATTA TGCCTAAAAA AACTGACAAA 14640  
GAATCTATCG ATGCAAAATAT TAAAAGCTTA ATACCTTTCC TTTGTTACCC TATAACAAAA 14700  
AAAGGAATTA AGACTTCATT GTCAAAATTG AAGAGTGTAG TTAGTGGAGA TATATTATCA 14760  
TATTCTATAG CTGGACGTAA TGAAGTATTC AGCAACAAGC TTATAAACCA CAAGCATATG 14820  
AATATCCTAA AATGGCTAGA TCATGTTTTA AACTTTAGAT CAGCTGAACT TAATTACAAT 14880

SUBSTITUTE SHEET (RULE 26)

- 320 -

CATTATATA TGATAGAGTC CACATATCCT TACTTAAGTG AATTGTTAAA CAGTTTAAACA 14940  
 ACCAATGAGC TCAAGAAGCT GATTAAAATA ACAGGTAGTG TACTATACAA CCTTCCCAAC 15000  
 GAACAGTAAC TTAAACATC ATTAACAAGT TTGATCAAAT TTAGATGCTA ACACATCATA 15060  
 ATATTATAGT TATTAAAAA TATATATGCA AACTTTTCAA TAATTTAGCA TATTGATTCC 15120  
 AAAGTTATCA TTTTGGTCTT AAGGGGTTGA ATAAAAATCT AAAACTAACA ATTATACATG 15180  
 TGCATTTACA ACACAACGAG ACATTAGTTT TTGACACTTT TTTTCTCGT 15229

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp  
 1 5 10 15  
 Ser Tyr Leu Lys Gly Val Ile Ser Phe Ser Glu Cys Asn Ala Leu Gly  
 20 25 30  
 Ser Tyr Leu Phe Asn Gly Pro Tyr Leu Lys Asn Asp Tyr Thr Asn Leu  
 35 40 45  
 Ile Ser Arg Gln Ser Pro Leu Leu Glu His Met Asn Leu Lys Lys Leu  
 50 55 60  
 Thr Ile Thr Gln Ser Leu Ile Ser Arg Tyr His Lys Gly Glu Leu Lys  
 65 70 75 80  
 Leu Glu Glu Pro Thr Tyr Phe Gln Ser Leu Leu Met Thr Tyr Lys Ser  
 85 90 95  
 Met Ser Ser Ser Glu Gln Ile Ala Thr Thr Asn Leu Leu Lys Lys Ile  
 100 105 110  
 Ile Arg Arg Ala Ile Glu Ile Ser Asp Val Lys Val Tyr Ala Ile Leu  
 115 120 125

SUBSTITUTE SHEET (RULE 26)



- 321 -

Asn Lys Leu Gly Leu Lys Glu Lys Asp Arg Val Lys Pro Asn Asn Asn  
 130 135 140  
 Ser Gly Asp Glu Asn Ser Val Leu Thr Thr Ile Ile Lys Asp Asp Ile  
 145 150 155 160  
 Leu Ser Ala Val Glu Asn Asn Gln Ser Tyr Thr Asn Ser Asp Lys Asn  
 165 170 175  
 His Ser Val Asn Gln Asn Ile Thr Ile Lys Thr Thr Leu Leu Lys Lys  
 180 185 190  
 Leu Met Cys Ser Met Gln His Pro Pro Ser Trp Leu Ile His Trp Phe  
 195 200 205  
 Asn Leu Tyr Thr Lys Leu Asn Asn Ile Leu Thr Gln Tyr Arg Ser Asn  
 210 215 220  
 Glu Val Lys Ser His Gly Phe Ile Leu Ile Asp Asn Gln Thr Leu Ser  
 225 230 235 240  
 Asp Phe Gln Phe Ile Leu Asn Gln Tyr Gly Cys Ile Val Tyr His Lys  
 245 250 255  
 Gly Leu Lys Lys Ile Thr Thr Thr Tyr Asn Gln Phe Leu Thr Trp  
 260 265 270  
 Lys Asp Ile Ser Leu Ser Arg Leu Asn Val Cys Leu Ile Thr Trp Ile  
 275 280 285  
 Ser Asn Cys Leu Asn Thr Leu Asn Lys Ser Leu Gly Leu Arg Cys Gly  
 290 295 300  
 Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile  
 305 310 315 320  
 Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu  
 325 330 335  
 Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe  
 340 345 350  
 Arg Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala  
 355 360 365  
 Ile Lys Ala Gln Lys Asn Leu Leu Ser Arg Val Cys His Thr Leu Leu  
 370 375 380  
 Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu  
 385 390 395 400  
 Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu

SUBSTITUTE SHEET (RULE 26)

- 322 -

405	410	415
Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro		
420	425	430
Met Val Asp Glu Arg Gln Ala Met Asp Ala Val Arg Ile Asn Cys Asn		
435	440	445
Glu Thr Lys Phe Tyr Leu Leu Ser Asn Leu Ser Thr Leu Arg Gly Ala		
450	455	460
Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp		
465	470	475
Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr		
485	490	495
Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Lys Asp		
500	505	510
Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro		
515	520	525
Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro		
530	535	540
Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser		
545	550	555
His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser		
565	570	575
Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe		
580	585	590
Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn		
595	600	605
Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser		
610	615	620
Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln		
625	630	635
Ile Leu Ala Glu Lys Met Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro		
645	650	655
Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu		
660	665	670
Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr		
675	680	685

SUBSTITUTE SHEET (RULE 26)

- 323 -

Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe  
 690 695 700  
 Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu  
 705 710 715 720  
 Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr  
 725 730 735  
 Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe  
 740 745 750  
 Ile Lys Asp His Val Val Asn Leu Asn Lys Val Asp Glu Gln Ser Gly  
 755 760 765  
 Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu  
 770 775 780  
 Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly  
 785 790 795 800  
 Lys Phe Ser Ile Thr Ala Leu Ile Asn Gly Asp Asn Gln Ser Ile Asp  
 805 810 815  
 Ile Ser Lys Pro Val Arg Leu Ile Glu Gly Gln Thr His Ala Gln Ala  
 820 825 830  
 Asp Tyr Leu Leu Ala Leu Asn Ser Leu Lys Leu Leu Tyr Lys Glu Tyr  
 835 840 845  
 Ala Gly Ile Gly His Lys Leu Lys Gly Thr Glu Thr Tyr Ile Ser Arg  
 850 855 860  
 Asp Met Gln Phe Met Ser Lys Thr Ile Gln His Asn Gly Val Tyr Tyr  
 865 870 875 880  
 Pro Ala Ser Ile Lys Lys Val Leu Arg Val Gly Pro Trp Ile Asn Thr  
 885 890 895  
 Ile Leu Asp Asp Phe Lys Val Ser Leu Glu Ser Ile Gly Ser Leu Thr  
 900 905 910  
 Gln Glu Leu Glu Tyr Arg Gly Glu Ser Leu Leu Cys Ser Leu Ile Phe  
 915 920 925  
 Arg Asn Ile Trp Leu Tyr Asn Gln Ile Ala Leu Gln Leu Arg Asn His  
 930 935 940  
 Ala Leu Cys His Asn Lys Leu Tyr Leu Asp Ile Leu Lys Val Leu Lys  
 945 950 955 960

SUBSTITUTE SHEET (RULE 26)

- 324 -

His Leu Lys Thr Phe Phe Asn Leu Asp Ser Ile Asp Met Ala Leu Thr  
 965 970 975

Leu Tyr Met Asn Leu Pro Met Leu Phe Gly Gly Gly Asp Pro Asn Leu  
 980 985 990

Leu Tyr Arg Ser Phe Tyr Arg Arg Thr Pro Asp Phe Leu Thr Glu Ala  
 995 1000 1005

Ile Val His Ser Val Phe Val Leu Ser Tyr Tyr Thr Gly His Asp Leu  
 1010 1015 1020

Gln Asp Lys Leu Gln Asp Leu Pro Asp Asp Arg Leu Asn Lys Phe Leu  
 1025 1030 1035 1040

Thr Cys Ile Ile Thr Phe Asp Lys Asn Pro Asn Ala Glu Phe Val Thr  
 1045 1050 1055

Leu Met Arg Asp Pro Gln Ala Leu Gly Ser Glu Arg Gln Ala Lys Ile  
 1060 1065 1070

Thr Ser Glu Ile Asn Arg Leu Ala Val Thr Glu Val Leu Ser Ile Ala  
 1075 1080 1085

Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu  
 1090 1095 1100

Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His  
 1105 1110 1115 1120

Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys  
 1125 1130 1135

Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu  
 1140 1145 1150

Lys Thr Ser Ala Ile Asp Ser Thr Asp Ile Asn Arg Ala Thr Asp Met  
 1155 1160 1165

Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys  
 1170 1175 1180

Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr  
 1185 1190 1195 1200

Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile  
 1205 1210 1215

Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asp Ile Lys Tyr  
 1220 1225 1230

Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val

SUBSTITUTE SHEET (RULE 26)

- 325 -

1235	1240	1245
Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly 1250	1255	1260
Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val 1265	1270	1275 1280
Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp 1285	1290	1295
Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu 1300	1305	1310
Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe 1315	1320	1325
Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser 1330	1335	1340
Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn 1345	1350	1355 1360
Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr 1365	1370	1375
Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly 1380	1385	1390
Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn 1395	1400	1405
Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro 1410	1415	1420
Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile 1425	1430	1435 1440
Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr 1445	1450	1455
Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile 1460	1465	1470
Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn 1475	1480	1485
Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile 1490	1495	1500
Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu 1505	1510	1515 1520

SUBSTITUTE SHEET (RULE 26)

- 326 -

Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn  
 1525 1530 1535  
 Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala  
 1540 1545 1550  
 Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu  
 1555 1560 1565  
 Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu  
 1570 1575 1580  
 Gln Lys Val Ile Lys Tyr Ile Ile Asn Gln Asp Thr Ser Leu His Arg  
 1585 1590 1595 1600  
 Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asn  
 1605 1610 1615  
 Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His  
 1620 1625 1630  
 Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met  
 1635 1640 1645  
 Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe  
 1650 1655 1660  
 Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe  
 1665 1670 1675 1680  
 Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser  
 1685 1690 1695  
 Glu Leu Glu Asn Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr  
 1700 1705 1710  
 Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys  
 1715 1720 1725  
 Pro Lys Phe Gly Ile Ser Gly Asn Thr Glu Ser Met Met Thr Ser Thr  
 1730 1735 1740  
 Phe Ser Asn Lys Thr His Ile Lys Ser Ser Ala Val Ile Thr Arg Phe  
 1745 1750 1755 1760  
 Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile  
 1765 1770 1775  
 Asp Arg Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu  
 1780 1785 1790

SUBSTITUTE SHEET (RULE 26)

- 327 -

Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser  
 1795 1800 1805  
 Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val  
 1810 1815 1820  
 Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp  
 1825 1830 1835 1840  
 Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala  
 1845 1850 1855  
 Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg  
 1860 1865 1870  
 Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile  
 1875 1880 1885  
 Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu  
 1890 1895 1900  
 Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser  
 1905 1910 1915 1920  
 Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp  
 1925 1930 1935  
 Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp  
 1940 1945 1950  
 Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys  
 1955 1960 1965  
 Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu  
 1970 1975 1980  
 Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu  
 1985 1990 1995 2000  
 Lys Gly Ser Glu Val Tyr Leu Val Leu Thr Ile Gly Pro Ala Asn Ile  
 2005 2010 2015  
 Leu Pro Val Phe Asn Val Val Gln Asn Ala Lys Leu Ile Leu Ser Arg  
 2020 2025 2030  
 Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp  
 2035 2040 2045  
 Ala Asn Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys  
 2050 2055 2060  
 Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Ser Gly

SUBSTITUTE SHEET (RULE 26)

- 328 -

2065	2070	2075	2080
Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn			
2085	2090	2095	
Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His			
2100	2105	2110	
Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met			
2115	2120	2125	
Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr			
2130	2135	2140	
Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr			
2145	2150	2155	2160
Asn Leu Pro Asn Glu Gln			
2165			

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

ACGGGAAAAA AATGCGTACT ACAAACTTGC ACATTCGAAA AAAATGGGGC AAATAAGAAC	60
TTGATAAGTG CTATTTAAGT CTAACCTTTT CAATCAGAAA TGGGGTGCAA TTCACTGAGC	120
ATGATAAAGG TTAGATTACA AAATTTATTT GACAATGACG AAGTAGCATT GTTAAAAATA	180
ACATGTTATA CTGATAAATT AATCTTCTG ACCAATGCAT TAGCCAAAGC AGCAATACAT	240
ACAATTAAAT TAAACGGCAT AGTTTTTATA CATGTTATAA CAAGCAGTGA AGTGTGCCCT	300
GATAACAATA TTGTAGTGAA ATCTAACTTT ACAACAATGC CAATACTACA AAATGGAGGA	360
TACATATGGG AATTGATTGA GTTGACACAC TGCTCTCAAT TAAACGGTTT AATGGATGAT	420
AATTGTGAAA TCAAATTTTC TAAAAGACTA AGTGACTCAG TAATGACTAA TTATATGAAT	480
CAAATATCTG ACTTACTTGG GCTTGATCTC AATTCATGAA TTATGTTTAG TCTAATTCAA	540

SUBSTITUTE SHEET (RULE 26)



- 329 -

TAGACATGTG TTTATTACCA TTTTAGTTAA TATAAAAACT CATCAAAGGG AAATGGGGCA	600
AATAAACTCA CCTAATCAAT CAAACCATGA GCACTACAAA TGACAACACT ACTATGCAAA	660
GATTGATGAT CACAGACATG AGACCCCTGT CAATGGATTG AATAATAACA TCTCTTACCA	720
AAGAAATCAT CACACACAAA TTCATATACT TGATAAACAA TGAATGTATT GTAAGAAAAC	780
TTGATGAAAG ACAAGCTACA TTTACATTCT TAGTCAATTA TGAGATGAAG CTACTGCACA	840
AAGTAGGGAG TACCAAATAC AAAAAATACA CTGAATATAA TACAAAATAT GGCACTTTCC	900
CCATGCCTAT ATTTATCAAT CACGGCGGGT TTCTAGAATG TATTGGCATT AAGCCTACAA	960
AACACACTCC TATAATATAC AAATATGACC TCAACCCGTG AATTCCAACA AAAAAACCAA	1020
CCCAACCAAA CCAAACCTATT CCTCAAACAA CAGTGCTCAA TAGTTAAGAA GGAGCTAATC	1080
CATTTTAGTA ATTA AAAATA AAAGTAAAGC CAATAACATA AATTGGGGCA AATACAAAGA	1140
TGGCTCTTAG CAAAGTCAAG TTGAATGATA CATTAAATAA GGATCAGCTG CTGTCATCCA	1200
GCAAATACAC TATTCACGT AGTACAGGAG ATAATATTGA CACTCCCAAT TATGATGTGC	1260
AAAAACACCT AAACAAACTA TGTGGTATGC TATTAATCAC TGAAGATGCA AATCATAAAT	1320
TCACAGGATT AATAGGTATG TTATATGCTA TGTCCAGGTT AGGAAGGGAA GACACTATAA	1380
AGATACTTAA AGATGCTGGA TATCATGTTA AAGCTAATGG AGTAGATATA ACAACATATC	1440
GTCAAGATAT AAATGGAAAG GAAATGAAAT TCGAAGTATT AACATTATCA AGCTTGACAT	1500
CAGAAATACA AGTCAATATT GAGATAGAAT CTAGAAAGTC CTACAAAAAA ATGCTAAAAG	1560
AGATGGGAGA AGTGGCTCCA GAATATAGGC ATGATTCTCC AGACTGTGGG ATGATAATAC	1620
TGTGTATAGC TGCACTTGTG ATAACCAAT TAGCAGCAGG AGACAGATCA GGTCTTACAG	1680
CAGTAATTAG GAGGGCAAAC AATGTCTTAA AAAACGAAAT AAAACGATAC AAGGGCCTCA	1740
TACCAAAGGA TATAGCTAAC AGTTTTTATG AAGTGTTTGA AAAACACCCT CATCTTATAG	1800
ATGTTTTCGT GCACTTTGGC ATTGCACAAT CATCCACAAG AGGGGGTAGT AGAGTTGAAG	1860
GAATCTTTGC AGGATTGTTT ATGAATGCCT ATGGTTCAGG GCAAGTAATG CTAAGATGGG	1920
GAGTTTTAGC CAAATCTGTA AAAAAATACA TGCTAGGACA TGCTAGTGTC CAGGCAGAAA	1980
TGGAGCAAGT TGTGGAAGTC TATGAGTATG CACAGAAGTT GGGAGGAGAA GCTGGATTCT	2040
ACCATATATT GAACAATCCA AAAGCATCAT TGCTGTCATT AACTCAATTT CCCAACTTCT	2100

SUBSTITUTE SHEET (RULE 26)

- 330 -

CAAGTGTGGT	CCTAGGCAAT	GCAGCAGGTC	TAGGCATAAT	GGGAGAGTAT	AGAGGTACAC	2160
CAAGAAACCA	GGATCTTTAT	GATGCAGCTA	AAGCATATGC	AGAGCAACTC	AAAGAAAATG	2220
GAGTAATAAA	CTACAGTGTA	TTAGACTTAA	CAGCAGAAGA	ATTGGAAGCC	ATAAAGCATC	2280
AACTCAACCC	CAAAGAAGAT	GATGTAGAGC	TTTAAGTTAA	CAAAAAATAC	GGGGCAAATA	2340
AGTCAACATG	GAGAAGTTTG	CACCTGAATT	TCATGGAGAA	GATGCAAATA	ACAAAGCTAC	2400
CAAATTCCTA	GAATCAATAA	AGGGCAAGTT	CGCATCATCC	AAAGATCCTA	AGAAGAAAGA	2460
TAGCATAATA	TCTGTAACT	CAATAGATAT	AGAAGTAACT	AAAGAGAGCC	CGATAACATC	2520
TGGCACCAAC	ATCATCAATC	CAACAAGTGA	AGCCGACAGT	ACCCCAAGAA	CAAAAGCCAA	2580
CTACCCAAGA	AAACCCCTAG	TAAGCTTCAA	AGAAGATCTC	ACCCCAAGTG	ACAACCCTTT	2640
TTCTAAGTTG	TACAAGGAAA	CAATAGAAAC	ATTTGATAAC	AATGAAGAAG	AATCTAGCTA	2700
CTCATATGAA	GAGATAAATG	ATCAAACAAA	TGACAACATT	ACAGCAAGAC	TAGATAGAAT	2760
TGATGAAAAA	TTAAGTGAAA	TATTAGGAAT	GCTCCATACA	TTAGTAGTTG	CAAGTGCAGG	2820
ACCCACTTCA	GCTCGCGATG	GAATAAGAGA	TGCTATGGTT	GGTCTAAGAG	AAGAGATGAT	2880
AGAAAAAATA	AGAGCGGAAG	CATTAATGAC	CAATGATAGG	TTAGAGGCTA	TGGCAAGACT	2940
TAGGAATGAG	GAAAGCGAAA	AAATGGCAAA	AGACACCTCA	GATGAAGTGT	CTCTTAATCC	3000
AACTTCCAAA	AAATTGAGTG	ACTTGTTGGA	AGACAACGAT	AGTGACAATG	ATCTATCACT	3060
TGATGATTTT	TGATCAGCGA	TCAACTCACT	CAGCAATCAA	CAACATCAAT	AAAACAGACA	3120
TCAATCCATT	GAATCAACTG	CCAGACCGAA	CAAACAAACG	TCCATCAGTA	GAACCACCAA	3180
CCAATCAATC	AACCAATTGA	TCAATCAGCA	ACCCGACAAA	ATTAACAATA	TAGTAACAAA	3240
AAAAGAACAA	GATGGGGCAA	ATATGGAAAC	ATACGTGAAC	AAGCTTCACG	AAGGCTCCAC	3300
ATACACAGCA	GCTGTTCACT	ACAATGTTCT	AGAAAAAGAT	GATGATCCTG	CATCACTAAC	3360
AATATGGGTG	CCTATGTTCC	AGTCATCTGT	GCCAGCAGAC	TTGCTCATAA	AAGAAGTTGC	3420
AAGCATCAAT	ATACTAGTGA	AGCAGATCTC	TACGCCCAAA	GGACCTTCAC	TACGAGTCAC	3480
GATTAACCTA	AGAAGTGCTG	TGCTGGCTCA	AATGCCTAGT	AATTTTCATCA	TAAGCGCAAA	3540
TGTATCATT	GATGAAAGAA	GCAAATTAGC	ATATGATGTA	ACTACACCTT	GTGAAATCAA	3600
AGCATGCAGT	CTAACATGCT	TAAAAGTAAA	AAGTATGTTA	ACTACAGTCA	AAGATCTTAC	3660

SUBSTITUTE SHEET (RULE 26)

- 331 -

CATGAAGACA TTCAACCCCA CTCATGAGAT CATTGCTCTA TGTGAATTG AAAATATTAT 3720  
GACATCAAAA AGAGTAATAA TACCAACCTA TCTAAGATCA ATTAGTGTCA AGAACAAGGA 3780  
TCTGAACTCA CTAGAAAATA TAGCAACCAC CGAATTCAAA AATGCTATCA CCAATGCAAA 3840  
AATTATTCCT TATGCAGGAT TAGTGTTAGT TATCACAGTT ACTGACAATA AAGGAGCATT 3900  
CAAATATATC AAACCACAGA GTCAATTTAT AGTAGATCTT GGTGCCTACC TAGAAAAAGA 3960  
GAGCATATAT TATGTGACTA CTAATTGGAA GCATACAGCT ACACGTTTTT CAATCAAACC 4020  
ACTAGAGGAT TAAACTTAAT TATCAACACT GAATGACAGG TCCACATATA TCCTCAAAC 4080  
ACACACTATA TCCAAACATC ATAAACATCT ACACTACACA CTTTCATCACA CAAACCAATC 4140  
CCACTCAAAA TCCAAAATCA CTACCAGCCA CTATCCGCTA GACCTAGAGT GCGAATAGGC 4200  
AAATAAAACC AAAATATGGG GTAAATAGAC ATTAGTTAGA GTTCAATCAA TCTTAACAAC 4260  
CATTTATACC GCCAATTCAA CACATATACT ATAAATCTTA AAATGGGAAA TACATCCATC 4320  
ACAATAGAAC TCACAAGCAA ATTTTGGCCC TATTTTACAC TAATACATAT GATCTTAACT 4380  
CTAATCTTTT TACTAATTAT AATCACTATC ATGATTGCAA CACTAAATAA GCTAAGTGAA 4440  
CACAAAGCAT TCTGCAACAA AACTCTTGAA CTAGGACAGA TGTACCAAAT CAACACACAG 4500  
AGTTCCACCA TTATGCTGTG TCAAACCATA ATCCTGTATA TACAAACAAA CAAATCCAAT 4560  
CCTCTCACAG AGTCACGGTG TCGCAAAACC ACGCTAACCA TCATGGTAGC ATAGAGTAGT 4620  
TATTTAAAAA TTAACATAAT GATGAATTGT TAGTATGAGA TCAAAAACAA CATGGGGCA 4680  
AATGCAACCA TGTCCAAACA CAAGAATCAA CGCACTGCCA GGACTCTAGA AAAGACCTGG 4740  
GATACTCTTA ATCATCTAAT TGTAATATCC TCTTGTTTAT ACAGATTAAA TTAAAAATCT 4800  
ATAGCACAAA TAGCACTATC AGTTTTGGCA ATGATAATCT CAACCTCTCT CATAATTGCA 4860  
GCCATAATAT TCATCATCTC TGCCAATCAC AAAGTTACAC TAACAACGGT CACAGTTCAA 4920  
ACAATAAAAA ACCCACTGA AAAAAACATC ACCACCTACC CTACTCAAGT CTCACCAGAA 4980  
AGGGTTAGTT CATCCAAGCA ACCCACAACC ACATCACCAA TCCACACAAG TTCAGCTACA 5040  
ACATCACCCA ATACAAAATC AGAAACACAC CATACAACAG CACAAACCAA AGGCAGAACC 5100  
ACCACTTCAA CACAGACCAA CAAGCCAAGC ACAAACCAC GTCCAAAAAA TCCACCAAAA 5160  
AAAGATGATT ACCATTTTGA AGTGTTCAAC TTCGTTCCCT GCAGTATATG TGGCAACAAT 5220

SUBSTITUTE SHEET (RULE 26)

- 332 -

CAACTTTGCA AATCCATCTG CAAAACAATA CCAAGCAACA AACCAAAGAA GAAACCAACC	5280
ATCAAACCCA CAAACAAACC AACCACCAAA ACCACAAACA AAAGAGACCC AAAAACACCA	5340
GCCAAAACGA CGAAAAAAGA AACTACCACC AACCACAAAC AAAAATAAC CCTCAAGACC	5400
ACAGAAAGAG ACACCAGCAC CTCACAATCC ACTGCACTCG ACACAACCAC ATTAATAACAC	5460
ACAGTCCAAC AGCAATCCCT CCTCTCAACC ACCCCCGAAA ACACACCCAA CTCCACACAA	5520
ACACCCACAG CATCCGAGCC CTCCACACCA AACTCCACCC AAAAACCCCA GCCACATGCT	5580
TAGTTATTCA AAAACTACAT CTTAGCAGAG AACCGTGATC TATCAAGCAA GAACGAAATT	5640
AAACCTGGGG CAAATAACCA TGGAGTTGAT GATCCACAAG TCAAGTGCAA TCTTCCTAAC	5700
TCTTGCTATT AATGCATTGT ACCTCACCTC AAGTCAGAAC ATAAGTGAGG AGTTTTACCA	5760
ATCGACATGT AGTGCAGTTA GCAGAGGTTA TTTTAGTGCT TTAAGAACAG GTTGGTATAC	5820
TAGTGTGATA ACAATAGAAT TAAGTAATAT AAAAGAAACC AAATGCAATG GAACTGACAC	5880
TAAAGTAAAA CTTATGAAAC AAGAATTAGA TAAGTATAAG AATGCAGTAA CAGAATTACA	5940
GCTACTTATG CAAAACACAC CAGCTGTCAA CAACCGGGCC AGAAGAGAAG CACCACAGTA	6000
TATGAACTAC ACAATCAATA CCACTAAAAA CCTAAATGTA TCAATAAGCA AGAAGAGGAA	6060
ACGAAGATTT CTAGGCTTCT TGTTAGGTGT GGGATCTGCA ATAGCAAGTG GTATAGCTGT	6120
ATCAAAAGTT CTACACCTTG AAGGAGAAGT GAACAAGATC AAAAATGCTT TGTTGTCTAC	6180
AAACAAAGCT GTAGTCAGTT TATCAAATGG GGTCAGTGTT TTAACCAGCA AAGTGTTAGA	6240
TCTCAAGAAT TACATAAATA ACCAATTATT ACCCATAGTA AATCAACAGA GCTGTGCGAT	6300
CTCCAACATT GAAACAGTTA TAGAATTCCA GCAGAAGAAC AGCAGATTGT TGGAAATCAC	6360
CAGAGAATTT AGTGTCAATG CAGGTGTAAC AACACCTTTA AGCACTTACA TGTGACAAA	6420
CAGTGAGTTA CTATCATTAA TCAATGATAT GCCTATAACA AATGATCAGA AAAAATTAAT	6480
GTCAAGCAAT GTTCAGATAG TAAGGCAACA AAGTTATTCC ATCATGTCTA TAATAAGGA	6540
AGAAGTCCTT GCATATGTTG TACAGCTGCC TATCTATGGT GTAATAGATA CACCTTGCTG	6600
GAAATTGCAC ACATCGCCTC TATGCACTAC CAACATCAAA GAAGGATCAA ATATTGTTT	6660
AACAAGGACT GATAGAGGAT GGTATTGTGA TAATGCAGGA TCAGTATCCT TCTTTCCACA	6720
GGCTGACACT TGTAAGTAC AGTCCAATCG AGTATTTTGT GACACTATGA ACAGTTTGAC	6780

SUBSTITUTE SHEET (RULE 26)

- 333 -

ATTACCAAGT GAAGTCAGCC TTTGTAACAC TGACATATTC AATCCAAGT ATGACTGCAA	6840
AATTATGACA TCAAAAACAG ACATAAGCAG CTCAGTAATT ACTTCTCTTG GAGCTATAGT	6900
GTCATGCTAT GGTAAACTA AATGCACTGC ATCCAACAAA AATCGTGGGA TTATAAAGAC	6960
ATTTTCTAAT GGTGTGACT ATGTGTCAAA CAAAGGAGTA GATACTGTGT CAGTGGGCAA	7020
CACTTTATAC TATGTAAACA AGCTGGAAGG CAAGAACCTT TATGTAAAAG GGAACCTAT	7080
AATAAATTAC TATGACCCTC TAGTGTTTCC TTCTGATGAG TTTGATGCAT CAATATCTCA	7140
AGTCAATGAA AAAATCAATC AAAGTTTAGC TTTTATTCGT AGATCTGATG AATTACTACA	7200
TAATGTAAAT ACTGGCAAAT CTACTACAAA TATTATGATA ACTACAATTA TTATAGTAAT	7260
CATTGTAGTA TTGTTATCAT TAATAGCTAT TGGTTTACTG TTGTATTGTA AAGCCAAAAA	7320
CACACCAGTT ACACTAAGCA AAGACCAACT AAGTGAATC AATAATATTG CATTACAGCA	7380
ATAGACAAAA AACCACCTGA TCATGTTTCA ACAACAATCT GCTGACCACC AATCCCAAAT	7440
CAACTTACAA CAAATATTTT AACATCACAG TACAGGCTGA ATCATTTCCT CACATCATGC	7500
TACCCACATA ACTAAGCTAG ATCCTTAACT TATAGTTACA TAAAAACCTC AAGTATCACA	7560
ATCAACCACT AAATCAACAC ATCATTACAA AAATTAACAG CTGGGGCAAA TATGTCGCGA	7620
AGAAATCCTT GTAAATTTGA GATTAGAGGT CATTGCTTGA ATGGTAGAAG ATGTCACTAC	7680
AGTCATAATT ACTTTGAATG GCCTCCTCAT GCATTACTAG TGAGGCAAAA CTTCATGTTA	7740
AACAAGATAC TCAAGTCAAT GGACAAAAGC ATAGACACTT TGTCTGAAAT AAGTGGAGCT	7800
GCTGAACTGG ATAGAACAGA AGAATATGCT CTTGGTATAG TTGGAGTGCT AGAGAGTTAC	7860
ATAGGATCTA TAAACAACAT AACAAAACAA TCAGCATGTG TTGCTATGAG TAACTTCTT	7920
ATTGAGATCA ATAGTGATGA CATTAAAAAG CTTAGAGATA ATGAAGAACC CAATTCACCT	7980
AAGATAAGAG TGTACAATAC TGTTATATCA TACATTGAGA GCAATAGAAA AAACAACAAG	8040
CAAACCATCC ATCTGCTCAA GAGACTACCA GCAGACGTGC TGAAGAAGAC AATAAAGAAC	8100
ACATTAGATA TCCACAAAAG CATAACCATA AGCAATCCAA AAGAGTCAAC TGTGAATGAT	8160
CAAAATGACC AAACCAAAAA TAATGATATT ACCGGATAAA TATCCTTGTA GTATATCATC	8220
CATATTGATC TCAAGTGAAA GCATGGTTGC TACATTCAAT CATAAAAAACA TATTACAATT	8280
TAACCATAAC TATTTGGATA ACCACCAGCG TTTATTAAAT CATATATTTG ATGAAATTCA	8340

SUBSTITUTE SHEET (RULE 26)

- 334 -

TTGGACACCT AAAAATTAT TAGATGCCAC TCAACAATTT CTCCAACATC TTAACATCCC	8400
TGAAGATATA TATACAGTAT ATATATTAGT GTCATAATGC TTGACCATAA CGACTCTATG	8460
TCATCCAACC ATAAAACTAT TTTGATAAGG TTATGGGACA AAATGGATCC CATTATTAAT	8520
GGAAACTCTG CTAATGTGTA TCTAACTGAT AGTTATTTAA AAGGTGTTAT CTCTTTTTCA	8580
GAGTGTAATG CTTTAGGGAG TTATCTTTTT AACGGCCCTT ATCTTAAAAA TGATTACACC	8640
AACTTAATTA GTAGACAAAG CCCACTACTA GAGCATATGA ATCTTAAAAA ACTAACTATA	8700
ACACAGTCAT TAATATCTAG ATATCATAAA GGTGAACTGA AATTAGAAGA ACCAACTTAT	8760
TTCCAGTCAT TACTTATGAC ATATAAAAGT ATGTCCTCGT CTGAACAAAT TGCTACAAC	8820
AACTTACTTA AAAAAATAAT ACGAAGAGCC ATAGAAATAA GTGATGTAAA GGTGTACGCC	8880
ATCTTGAATA AACTAGGATT AAAGGAAAAG GACAGAGTTA AGCCCAACAA TAATTCAGGT	8940
GATGAAAAC CAGTACTTAC AACCATAATT AAAGATGATA TACTTTCGGC TGTGAAAAAC	9000
AATCAATCAT ATACAAATTC AGACAAAAGT CACTCAGTAA ATCAAAATAT CACTATCAAA	9060
ACAACACTCT TGAAAAAATT GATGTGTTCA ATGCAACATC CTCCATCATG GTTAATACAC	9120
TGGTTCAATT TATATACAAA ATTAAATAAC ATATTAACAC AATATCGATC AAATGAGGTA	9180
AAAAGTCATG GGTTTATATT AATAGATAAT CAACTTTTAA GTGGTTTTCA GTTTATTTTA	9240
AATCAATATG GTTGATCGT TTATCATAAA GGACTCAAAA AAATCACAAC TACTACTTAC	9300
AATCAATTTT TGACATGGAA AGACATCAGC CTTAGCAGAT TAAATGTTTG CTTAATTACT	9360
TGGATAAGTA ATTGTTTAAA TACATTAAAC AAAAGCTTAG GGCTGAGATG TGGATTCAAT	9420
AATGTTGTGT TATCACAATT ATTTCTTTAT GGAGATTGTA TACTGAAATT ATTTCATAAT	9480
GAAGGCTTCT ACATAATAAA AGAAGTAGAG GGATTTATTA TGTCTTTAAT TCTAAACATA	9540
ACAGAAGAAG ATCAATTTAA GAAACGATTT TATAATAGCA TGCTAAATAA CATCACAGAT	9600
GCAGCTATTA AGGCTCAAAA GGACCTACTA TCAAGAGTAT GTCACACTTT ATTAGACAAG	9660
ACAGTGTCTG ATAATATCAT AAATGGTAAA TGGATAATCC TATTAAGTAA ATTTCTTAAA	9720
TTGATTAAGC TTGCAGGTGA TAATAATCTC AATAACTTGA GTGAGCTATA TTTTCTCTTC	9780
AGAACTTTTG GACATCCAAT GGTGATGAA AGACAAGCAA TGGATTCTGT AAGAATTAAC	9840
TGTAATGAAA CTAGGTTCTA CTTATTAAGT AGTCTAAGTA CATTAAGAGG TGCTTTCATT	9900

SUBSTITUTE SHEET (RULE 26)

- 335 -

TATAGAATCA TAAAAGGGTT TGTAATACC TACAACAGAT GGCCACCTT AAGGAATGCT	9960
ATTGTCTAC CTCTAAGATG GTTAACTAC TATAACTTA ATACTTATCC ATCTCTACTT	10020
GAAATCACAG AAAATGATTT GATTATTTTA TCAGGATTGC GGTTCTATCG TGAGTTTCAT	10080
CTGCCTAAAA AAGTGGATCT TGAAATGATA ATAAATGACA AAGCCATTTC ACCTCCAAAA	10140
GATCTAATAT GGACTAGTTT TCCTAGAAAT TACATGCCAT CACATATACA AAATTATATA	10200
GAACATGAAA AGTTGAAGTT CTCTGAAAGC GACAGATCGA GAAGAGTACT AGAGTATTAC	10260
TTGAGAGATA ATAAATTCAA TGAATGCGAT CTATACAATT GTGTAGTCAA TCAAAGCTAT	10320
CTCAACAACCT CTAATCACGT GGTATCACTA ACTGGTAAAG AAAGAGAGCT CAGTGTAGGT	10380
AGAATGTTTG CTATGCAACC AGGTATGTTT AGGCAATCC AAATCTTAGC AGAGAAAATG	10440
ATAGCTGAAA ATATTTTACA ATTCTTCCCT GAGAGTTTGA CAAGATATGG TGATCTAGAG	10500
CTTCAAAGA TATTAGAATT AAAAGCAGGA ATAAGCAACA AGTCAAATCG TTATAATGAT	10560
AACTACAACA ATTATATCAG TAAATGTTCT ATCATTACAG ATCTTAGCAA ATTCAATCAG	10620
GCATTTAGAT ATGAAACATC ATGTATCTGC AGTGATGTAT TAGATGAACT GCATGGAGTA	10680
CAATCTCTGT TCTCTGGTT GCATTTAACA ATACCTCTTG TCACAATAAT ATGTACATAT	10740
AGACATGCAC CTCCTTTCAT AAAGGATCAT GTTGTTAATC TTAATGAGGT TGATGAACAA	10800
AGTGGATTAT ACAGATATCA TATGGGTGGT ATTGAGGGCT GGTGTCAAAA ACTGTGGACC	10860
ATTGAAGCTA TATCATTATT AGATCTAATA TCTCTCAAAG GGAAATTCTC TATCACAGCT	10920
CTGATAAATG GTGATAATCA GTCAATTGAT ATAAGCAAAC CAGTTAGACT TATAGAGGGT	10980
CAGACCCATG CACAAGCAGA TTATTTGTTA GCATTAAATA GCCTTAAATT GTTATATAAA	11040
GAGTATGCAG GTATAGGCCA TAAGCTTAAG GGAACAGAGA CCTATATATC CCGAGATATG	11100
CAGTTCATGA GCAAAACAAT CCAGCACAAT GGAGTGTACT ATCCAGCCAG TATCAAAAAA	11160
GTCCTGAGAG TAGGTCCATG GATAACACG ATACTTGATG ATTTTAAAGT TAGTTTAGAA	11220
TCTATAGGCA GCTTAACACA GGAGTTAGAA TACAGAGGAG AAAGCTTATT ATGCAGTTTA	11280
ATATTTAGGA ACATTGTTT ATACAATCAA ATTGCTTTC AACTCCGAAA TCATGCATTA	11340
TGTAACAATA AGCTATATTT AGATATATTG AAAGTATTAA AACACTTAAA AACTTTTTTT	11400
AATCTTGATA GCATTGATAT GGCTTTATCA TTGTATATGA ATTTGCCTAT GCTGTTTGGT	11460

SUBSTITUTE SHEET (RULE 26)

- 336 -

GGTGGTGATC CTAATTTGTT ATATCGAAGC TTTTATAGGA GAACTCCAGA CTCCTTACA 11520  
GAAGCTATAG TACATTCAGT GTTTGTGTTG AGCTATTATA CTGGTCACGA TTTACAAGAT 11580  
AAGCTCCAGG ATCTTCCAGA TGATAGACTG AACAAATTCT TGACATGTGT CATCACATTT 11640  
GATAAAAATC CCAATGCCGA GTTTGTAACA TTGATGAGGG ATCCACAGGC TTTAGGGTCT 11700  
GAAAGGCAAG CTAAAATTAC TAGTGAGATT AATAGATTAG CAGTAACAGA AGTCTTAAGT 11760  
ATAGCCCCAA ACAAATATT TTCTAAAAGT GCACAACATT ATACTACCAC TGAGATTGAT 11820  
CTAAATGACA TTATGCAAAA TATAGAACCA ACTTACCCTC ATGGATTAAG AGTTGTTTAT 11880  
GAAAGTTTAC CTTTTTATAA AGCAGAAAAA ATAGTTAATC TTATATCAGG AACAAAATCC 11940  
ATAACTAATA TACTTGAAAA AACATCAGCA ATAGATACAA CTGATATTAA TAGGGCTACT 12000  
GATATGATGA GGAAAAATAT AACTTTACTT ATAAGGATAC TTCCACTAGA TTGTAACAAA 12060  
GACAAAAGAG AGTTATTAAG TTTAGAAAAT CTTAGTATAA CTGAATTAAG CAAGTATGTA 12120  
AGAGAAAGAT CTGGTCATT ATCCAATATA GTAGGAGTAA CATCGCCAAG TATTATGTTG 12180  
ACAATGAACA TTAAATATAC AACTAGCACT ATAGCCAGTG GTATAATAAT AGAAAAATAT 12240  
AATGTTAATA GTTTAACTCG TGGTGAAAGA GGACCCACCA AGCCATGGGT AGGCTCATCC 12300  
ACGCAGGAGA AAAAAACAAT GCCAGTGATC AACAGACAAG TTTTAACCAA AAAGCAAAGA 12360  
GACCAAATAG ATTTATTAGC AAAATTAGAC TGGGTATATG CATCCATAGA CAACAAAGAT 12420  
GAATTCATGG AAGAACTGAG TACTGGAACA CTTGGACTGT CATATGAAAA AGCCAAAAAG 12480  
TTGTTTCCAC AATATCTAAG TGTCAATTAT TTACACCGTT TAACAGTCAG TAGTAGACCA 12540  
TGTGAATTCC CTGCATCAAT ACCAGCTTAT AGAACAACAA ATTATCATTT TGATACTAGT 12600  
CCTATCAATC ATGTATTAAC AGAAAAGTAT GGAGATGAAG ATATCGACAT TGTGTTTCAA 12660  
AATTGCATAA GTTTTGGTCT TAGCCTGATG TCGGTTGTGG AACAAATTCAC AAACATATGT 12720  
CCTAATAGAA TTATTCTCAT ACCGAAGCTG AATGAGATAC ATTTGATGAA ACCTCCTATA 12780  
TTTACAGGAG ATGTTGATAT CATCAAGTTG AAGCAAGTGA TACAAAAGCA GCACATGTTG 12840  
CTACCAGATA AAATAAGTTT AACCCAATAT GTAGAATTAT TCTTAAGTAA CAAAGCACTT 12900  
AAATCTGGAT CTCACATCAA CTCTAATTTA ATATTAGTAC ATAAAATGTC TGATTATTTT 12960  
CATAATGCTT ATATTTTAAG TACTAATTTA GCTGGACATT GGATTCTGAT TATTCAACTT 13020

SUBSTITUTE SHEET (RULE 26)



- 337 -

ATGAAAGATT CAAAAGGTAT TTTTGAAAAA GATTGGGGAG AGGGGTACAT AACTGATCAT 13080  
ATGTTTCATTA ATTTGAATGT TTTCTTTAAT GCTTATAAGA CTTATTTGCT ATGTTTTTCAT 13140  
AAAGGTTATG GTAAAGCAAA ATTAGAATGT GATATGAACA CTTGAGATCT TCTTTGTGTT 13200  
TTGGAGTTAA TAGACAGTAG CTAAGGAAA TCTATGTCTA AAGTTTTCCT AGAACAAAAA 13260  
GTCATAAAAT ACATAGTCAA TCAAGACACA AGTTTGCCTA GAATAAAAGG CTGTCACAGT 13320  
TTTAAGTTGT GGTTTTTTAAA ACGCCTTAAT AATGCTAAAT TTACCGTATG CCCTTGGGTT 13380  
GTTAACATAG ATTATCACCC AACACACATG AAAGCTATAT TATCTTACAT AGATTTAGTT 13440  
AGAATGGGGT TAATAAATGT AGATAAATTA ACCATTAAAA ATAAAAACAA ATTCAATGAT 13500  
GAATTTTACA CATCAAATCT CTTTACATT AGTTATAACT TTTCAGACAA CACTCATTG 13560  
CTAACAAAAC AAATAAGAAT TGCTAATTCA GAATTAGAAG ATAATTATAA CAAACTATAT 13620  
CACCCAACCC CAGAACTTT AGAAAATATG TCATTAATTC CTGTTAAAAG TAATAATAGT 13680  
AACAAACCTA AATTTTGTAT AAGTGGAAT ACCGAATCTA TGATGATGTC AACATTCTCT 13740  
AGTAAATGC ATATTAAATC TTCCACTGTT ACCACAAGAT TCAATTATAG CAAACAAGAC 13800  
TTGTACAATT TATTTCCAAT TGTTGTGATA GACAAGATTA TAGATCATTC AGGTAATACA 13860  
GCAAAATCTA ACCAACTTA CACCACCACT TCACATCAGA CATCTTTAGT AAGGAATAGT 13920  
GCATCACTTT ATTGCATGCT TCCTTGGCAT CATGTCAATA GATTTAACTT TGTATTTAGT 13980  
TCCACAGGAT GCAAGATCAG TATAGAGTAT ATTTTAAAAG ATCTTAAGAT TAAGGACCCC 14040  
AGTTGTATAG CATTCATAGG TGAAGGAGCT GGTAACTTAT TATTACGTAC GGTAGTAGAA 14100  
CTTCATCCAG ACATAAGATA CATTACAGA AGTTTAAAAG ATTGCAATGA TCATAGTTTA 14160  
CCTATTGAAT TTCTAAGGTT ATACAACGGG CATATAAACA TAGATTATGG TGAGAATTTA 14220  
ACCATTCTTG CTACAGATGC AACTAATAAC ATTCATTGGT CTTATTTACA TATAAAATTT 14280  
GCAGAACCTA TTAGCATCTT TGTCTGCGAT GCTGAATTAC CTGTTACAGC CAATTGGAGT 14340  
AAAATTATAA TTGAATGGAG TAAGCATGTA AGAAAGTGCA AGTACTGTTT TTCTGTAAAT 14400  
AGATGCATTT TAATTGCAA ATATCATGCT CAAGATGACA TTGATTTCAA ATTAGATAAC 14460  
ATTACTATAT TAAAACTTA CGTGTGCCTA GGTAGCAAGT TAAAAGGATC TGAAGTTTAC 14520  
TTAATCCTTA CAATAGGCCC TGCAAATATA CTCCTGTTT TTGATGTTGT ACAAATGCT 14580

SUBSTITUTE SHEET (RULE 26)

- 338 -

AAATTGATAC TTTCAAGAAC TAAAAATTC ATTATGCCTA AAAAACTGA CAAGGAATCT 14640  
 ATCGATGCAA ATATTAAAAG CTTAATACCT TTCCTTTGTT ACCCTATAAC AAAAAAGGA 14700  
 ATTAAGACTT CATTGTCAAA ATTGAAGAGT GTAGTTAATG GAGATATATT ATCATATTCT 14760  
 ATAGCTGGAC GTAATGAAGT ATTCAGCAAC AAGCTTATAA ACCACAAGCA TATGAATATC 14820  
 CTAAGTGGC TAGATCATGT TTTAAATTTT AGATCAGCTG AACTTAATTA CAATCATTTA 14880  
 TACATGATAG AGTCCACATA TCCTTACTTA AGTGAATTGT TAAATAGTTT AACAACCAAT 14940  
 GAGCTCAAGA AGCTGATTAA AATAACAGGT AGTGTGCTAT ACAACCTTCC CAACGAACAG 15000  
 TAGTTTAAAA TATCATTAAC AAGTTTGGTC AAATTTAGAT GCTAACACAT CATTATATTA 15060  
 TAGTTATTAA AGAATATACA AACTTTTCAA TAATTTAGCA TATTGATTCC AAAATTATCA 15120  
 TTTTAGTCTT AAGGGGTAA ATAAAAGTCT AAAACTAACA ATTATACATG TGCATTACAA 15180  
 ACACAACGAG ACATTAGTTT TTGACACTTT TTTTCTCGT 15219

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met	Asp	Pro	Ile	Ile	Asn	Gly	Asn	Ser	Ala	Asn	Val	Tyr	Leu	Thr	Asp
1				5					10					15	
Ser	Tyr	Leu	Lys	Gly	Val	Ile	Ser	Phe	Ser	Glu	Cys	Asn	Ala	Leu	Gly
			20					25					30		
Ser	Tyr	Leu	Phe	Asn	Gly	Pro	Tyr	Leu	Lys	Asn	Asp	Tyr	Thr	Asn	Leu
		35					40					45			
Ile	Ser	Arg	Gln	Ser	Pro	Leu	Leu	Glu	His	Met	Asn	Leu	Lys	Lys	Leu
		50				55					60				
Thr	Ile	Thr	Gln	Ser	Leu	Ile	Ser	Arg	Tyr	His	Lys	Gly	Glu	Leu	Lys
65					70					75				80	

SUBSTITUTE SHEET (RULE 26)

- 339 -

Leu Glu Glu Pro Thr Tyr Phe Gln Ser Leu Leu Met Thr Tyr Lys Ser  
 85 90 95

Met Ser Ser Ser Glu Gln Ile Ala Thr Thr Asn Leu Leu Lys Lys Ile  
 100 105 110

Ile Arg Arg Ala Ile Glu Ile Ser Asp Val Lys Val Tyr Ala Ile Leu  
 115 120 125

Asn Lys Leu Gly Leu Lys Glu Lys Asp Arg Val Lys Pro Asn Asn Asn  
 130 135 140

Ser Gly Asp Glu Asn Ser Val Leu Thr Thr Ile Ile Lys Asp Asp Ile  
 145 150 155 160

Leu Ser Ala Val Glu Asn Asn Gln Ser Tyr Thr Asn Ser Asp Lys Ser  
 165 170 175

His Ser Val Asn Gln Asn Ile Thr Ile Lys Thr Thr Leu Leu Lys Lys  
 180 185 190

Leu Met Cys Ser Met Gln His Pro Pro Ser Trp Leu Ile His Trp Phe  
 195 200 205

Asn Leu Tyr Thr Lys Leu Asn Asn Ile Leu Thr Gln Tyr Arg Ser Asn  
 210 215 220

Glu Val Lys Ser His Gly Phe Ile Leu Ile Asp Asn Gln Thr Leu Ser  
 225 230 235 240

Gly Phe Gln Phe Ile Leu Asn Gln Tyr Gly Cys Ile Val Tyr His Lys  
 245 250 255

Gly Leu Lys Lys Ile Thr Thr Thr Tyr Asn Gln Phe Leu Thr Trp  
 260 265 270

Lys Asp Ile Ser Leu Ser Arg Leu Asn Val Cys Leu Ile Thr Trp Ile  
 275 280 285

Ser Asn Cys Leu Asn Thr Leu Asn Lys Ser Leu Gly Leu Arg Cys Gly  
 290 295 300

Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile  
 305 310 315 320

Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu  
 325 330 335

Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe  
 340 345 350

SUBSTITUTE SHEET (RULE 26)

- 340 -

Lys Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala  
 355 360 365  
 Ile Lys Ala Gln Lys Asp Leu Leu Ser Arg Val Cys His Thr Leu Leu  
 370 375 380  
 Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu  
 385 390 395 400  
 Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu  
 405 410 415  
 Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro  
 420 425 430  
 Met Val Asp Glu Arg Gln Ala Met Asp Ser Val Arg Ile Asn Cys Asn  
 435 440 445  
 Glu Thr Arg Phe Tyr Leu Leu Ser Ser Leu Ser Thr Leu Arg Gly Ala  
 450 455 460  
 Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp  
 465 470 475 480  
 Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr  
 485 490 495  
 Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Asn Asp  
 500 505 510  
 Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro  
 515 520 525  
 Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro  
 530 535 540  
 Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser  
 545 550 555 560  
 His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser  
 565 570 575  
 Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe  
 580 585 590  
 Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn  
 595 600 605  
 Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser  
 610 615 620  
 Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln

SUBSTITUTE SHEET (RULE 26)

- 341 -

625		630		635		640
Ile Leu Ala Glu Lys Met	Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro					
	645		650		655	
Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu						
	660		665		670	
Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr						
	675		680		685	
Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe						
	690		695		700	
Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu						
	705		710		715	720
Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr						
	725		730		735	
Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe						
	740		745		750	
Ile Lys Asp His Val Val Asn Leu Asn Glu Val Asp Glu Gln Ser Gly						
	755		760		765	
Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu						
	770		775		780	
Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly						
	785		790		795	800
Lys Phe Ser Ile Thr Ala Leu Ile Asn Gly Asp Asn Gln Ser Ile Asp						
	805		810		815	
Ile Ser Lys Pro Val Arg Leu Ile Glu Gly Gln Thr His Ala Gln Ala						
	820		825		830	
Asp Tyr Leu Leu Ala Leu Asn Ser Leu Lys Leu Leu Tyr Lys Glu Tyr						
	835		840		845	
Ala Gly Ile Gly His Lys Leu Lys Gly Thr Glu Thr Tyr Ile Ser Arg						
	850		855		860	
Asp Met Gln Phe Met Ser Lys Thr Ile Gln His Asn Gly Val Tyr Tyr						
	865		870		875	880
Pro Ala Ser Ile Lys Lys Val Leu Arg Val Gly Pro Trp Ile Asn Thr						
	885		890		895	
Ile Leu Asp Asp Phe Lys Val Ser Leu Glu Ser Ile Gly Ser Leu Thr						
	900		905		910	

SUBSTITUTE SHEET (RULE 26)

- 342 -

Gln Glu Leu Glu Tyr Arg Gly Glu Ser Leu Leu Cys Ser Leu Ile Phe  
 915 920 925  
 Arg Asn Ile Trp Leu Tyr Asn Gln Ile Ala Leu Gln Leu Arg Asn His  
 930 935 940  
 Ala Leu Cys Asn Asn Lys Leu Tyr Leu Asp Ile Leu Lys Val Leu Lys  
 945 950 955 960  
 His Leu Lys Thr Phe Phe Asn Leu Asp Ser Ile Asp Met Ala Leu Ser  
 965 970 975  
 Leu Tyr Met Asn Leu Pro Met Leu Phe Gly Gly Gly Asp Pro Asn Leu  
 980 985 990  
 Leu Tyr Arg Ser Phe Tyr Arg Arg Thr Pro Asp Phe Leu Thr Glu Ala  
 995 1000 1005  
 Ile Val His Ser Val Phe Val Leu Ser Tyr Tyr Thr Gly His Asp Leu  
 1010 1015 1020  
 Gln Asp Lys Leu Gln Asp Leu Pro Asp Asp Arg Leu Asn Lys Phe Leu  
 1025 1030 1035 1040  
 Thr Cys Val Ile Thr Phe Asp Lys Asn Pro Asn Ala Glu Phe Val Thr  
 1045 1050 1055  
 Leu Met Arg Asp Pro Gln Ala Leu Gly Ser Glu Arg Gln Ala Lys Ile  
 1060 1065 1070  
 Thr Ser Glu Ile Asn Arg Leu Ala Val Thr Glu Val Leu Ser Ile Ala  
 1075 1080 1085  
 Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu  
 1090 1095 1100  
 Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His  
 1105 1110 1115 1120  
 Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys  
 1125 1130 1135  
 Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu  
 1140 1145 1150  
 Lys Thr Ser Ala Ile Asp Thr Thr Asp Ile Asn Arg Ala Thr Asp Met  
 1155 1160 1165  
 Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys  
 1170 1175 1180

SUBSTITUTE SHEET (RULE 26)

- 343 -

Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr  
 1185 1190 1195 1200  
 Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile  
 1205 1210 1215  
 Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asn Ile Lys Tyr  
 1220 1225 1230  
 Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val  
 1235 1240 1245  
 Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly  
 1250 1255 1260  
 Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val  
 1265 1270 1275 1280  
 Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp  
 1285 1290 1295  
 Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu  
 1300 1305 1310  
 Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe  
 1315 1320 1325  
 Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser  
 1330 1335 1340  
 Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn  
 1345 1350 1355 1360  
 Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr  
 1365 1370 1375  
 Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly  
 1380 1385 1390  
 Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn  
 1395 1400 1405  
 Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro  
 1410 1415 1420  
 Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile  
 1425 1430 1435 1440  
 Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr  
 1445 1450 1455  
 Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile

SUBSTITUTE SHEET (RULE 26)

- 344 -

1460	1465	1470
Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn 1475	1480	1485
Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile 1490	1495	1500
Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu 1505	1510	1515 1520
Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn 1525	1530	1535
Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala 1540	1545	1550
Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu 1555	1560	1565
Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu 1570	1575	1580
Gln Lys Val Ile Lys Tyr Ile Val Asn Gln Asp Thr Ser Leu Arg Arg 1585	1590	1595 1600
Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asn 1605	1610	1615
Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His 1620	1625	1630
Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met 1635	1640	1645
Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe 1650	1655	1660
Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe 1665	1670	1675 1680
Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser 1685	1690	1695
Glu Leu Glu Asp Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr 1700	1705	1710
Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys 1715	1720	1725
Pro Lys Phe Cys Ile Ser Gly Asn Thr Glu Ser Met Met Met Ser Thr 1730	1735	1740

SUBSTITUTE SHEET (RULE 26)



- 345 -

Phe Ser Ser Lys Met His Ile Lys Ser Ser Thr Val Thr Thr Arg Phe  
 1745 1750 1755 1760  
 Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile  
 1765 1770 1775  
 Asp Lys Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu  
 1780 1785 1790  
 Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser  
 1795 1800 1805  
 Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val  
 1810 1815 1820  
 Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp  
 1825 1830 1835 1840  
 Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala  
 1845 1850 1855  
 Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg  
 1860 1865 1870  
 Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile  
 1875 1880 1885  
 Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu  
 1890 1895 1900  
 Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser  
 1905 1910 1915 1920  
 Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp  
 1925 1930 1935  
 Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp  
 1940 1945 1950  
 Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys  
 1955 1960 1965  
 Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu  
 1970 1975 1980  
 Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu  
 1985 1990 1995 2000  
 Lys Gly Ser Glu Val Tyr Leu Ile Leu Thr Ile Gly Pro Ala Asn Ile  
 2005 2010 2015,

SUBSTITUTE SHEET (RULE 26)

- 346 -

Leu Pro Val Phe Asp Val Val Gln Asn Ala Lys Leu Ile Leu Ser Arg  
 2020 2025 2030  
 Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp  
 2035 2040 2045  
 Ala Asn Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys  
 2050 2055 2060  
 Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Asn Gly  
 2065 2070 2075 2080  
 Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn  
 2085 2090 2095  
 Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His  
 2100 2105 2110  
 Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met  
 2115 2120 2125  
 Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr  
 2130 2135 2140  
 Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr  
 2145 2150 2155 2160  
 Asn Leu Pro Asn Glu Gln  
 2165

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ACGGGAAAAA AATGCGTACT ACAAACTTGC ACATTCGAAA AAAATGGGGC AAATAAGAAC	60
TTGATAAGTG CTATTTAAGT CTAACCTTTT CAATCAGAAA TGGGGTGCAA TTCACTGAGC	120
ATGATAAAGG TTAGATTACA AAATTTATTT GACAATGACG AAGTAGCATT GTTAAAAATA	180
ACATGTTATA CTGATAAATT AATTCTTCTG ACCAATGCAT TAGCCAAAGC AGCAATACAT	240

SUBSTITUTE SHEET (RULE 26)

- 347 -

ACAATTAAAT TAAACGGCAT AGTTTTTATA CATGTTATAA CAAGCAGTGA AGTGTGCCCT	300
GATAACAATA TTGTAGTGAA ATCTAACTTT ACAACAATGC CAATACTACA AAATGGAGGA	360
TACATATGGG AATTGATTGA GTTGACACAC TGCTCTCAAT TAAACGGTTT AATGGATGAT	420
AATTGTGAAA TCAAATTTTC TAAAAGACTA AGTGACTCAG TAATGACTAA TTATATGAAT	480
CAAATATCTG ACTTACTTGG GCTTGATCTC AATTCATGAA TTATGTTTAG TCTAATTCAA	540
TAGACATGTG TTTATTACCA TTTTAGTTAA TATAAAAACT CATCAAAGGG AAATGGGGCA	600
AATAAACTCA CCTAATCAAT CAAACCATGA GCACTACAAA TGACAACACT ACTATGCAAA	660
GATTGATGAT CACAGACATG AGACCCCTGT CAATGGATTG AATAATAACA TCTCTTACCA	720
AAGAAATCAT CACACACAAA TTCATATACT TGATAAACAA TGAATGTATT GTAAGAAAAC	780
TTGATGAAAG ACAAGCTACA TTTACATTCT TAGTCAATTA TGAGATGAAG CTAAGCACA	840
AAGTAGGGAG TACCAAATAC AAAAAATACA CTGAATATAA TACAAAATAT GGCACCTTCC	900
CCATGCCTAT ATTTATCAAT CACGGCGGGT TTCTAGAATG TATTGGCATT AAGCCTACAA	960
AACACACTCC TATAATATAC AAATATGACC TCAACCCGTG AATTCCAACA AAAAAACCAA	1020
CCCAACCAA CCAAATATT CCTCAAACA CAGTGCTCAA TAGTTAAGAA GGAGCTAATC	1080
CATTTTAGTA ATTAATAATA AAAGTAAAGC CAATAACATA AATTGGGGCA AATACAAAGA	1140
TGGCTCTTAG CAAAGTCAAG TTGAATGATA CATTAAATAA GGATCAGCTG CTGTCATCCA	1200
GCAAATACAC TATTCAACGT AGTACAGGAG ATAATATTGA CACTCCCAAT TATGATGTGC	1260
AAAAACACCT AAACAACTA TGTGGTATGC TATTAATCAC TGAAGATGCA AATCATAAAT	1320
TCACAGGATT AATAGGTATG TTATATGCTA TGTCCAGGTT AGGAAGGGAA GACACTATAA	1380
AGATACTTAA AGATGCTGGA TATCATGTTA AAGCTAATGG AGTAGATATA ACAACATATC	1440
GTCAAGATAT AAATGGAAAG GAAATGAAAT TCGAAGTATT AACATTATCA AGCTTGACAT	1500
CAGAAATACA AGTCAATATT GAGATAGAAT CTAGAAAGTC CTACAAAAAA ATGCTAAAAG	1560
AGATGGGAGA AGTGGCTCCA GAATATAGGC ATGATTCTCC AGACTGTGGG ATGATAATAC	1620
TGTGTATAGC TGCACCTGTG ATAACCAAAT TAGCAGCAGG AGACAGATCA GGTCTTACAG	1680
CAGTAATTAG GAGGGCAAAC AATGTCTTAA AAAACGAAAT AAAACGATAC AAGGGCCTCA	1740
TACCAAAGGA TATAGCTAAC AGTTTTTATG AAGTGTTTGA AAAACACCCT CATCTTATAG	1800

SUBSTITUTE SHEET (RULE 26)

- 348 -

ATGTTTTTCGT	GCACTTTTGGC	ATTGCACAAT	CATCCACAAG	AGGGGGTAGT	AGAGTTGAAG	1860
GAATCTTTGC	AGGATTGTTT	ATGAATGCCT	ATGGTTCAGG	GCAAGTAATG	CTAAGATGGG	1920
GAGTTTTTAGC	CAAATCTGTA	AAAAATATCA	TGCTAGGACA	TGCTAGTGTC	CAGGCAGAAA	1980
TGGAGCAAGT	TGTGGAAGTC	TATGAGTATG	CACAGAAGTT	GGGAGGAGAA	GCTGGATTCT	2040
ACCATATATT	GAACAATCCA	AAAGCATCAT	TGCTGTCATT	AACTCAATTT	CCCAACTTCT	2100
CAAGTGTGGT	CCTAGGCAAT	GCAGCAGGTC	TAGGCATAAT	GGGAGAGTAT	AGAGGTACAC	2160
CAAGAAACCA	GGATCTTTAT	GATGCAGCTA	AAGCATATGC	AGAGCAACTC	AAAGAAAATG	2220
GAGTAATAAA	CTACAGTGTA	TTAGACTTAA	CAGCAGAAGA	ATTGGAAGCC	ATAAAGCATC	2280
AACTCAACCC	CAAAGAAGAT	GATGTAGAGC	TTTAAGTTAA	CAAAAAATAC	GGGGCAAATA	2340
AGTCAACATG	GAGAAGTTTG	CACCTGAATT	TCATGGAGAA	GATGCAAATA	ACAAAGCTAC	2400
CAAATTCCTA	GAATCAATAA	AGGGCAAGTT	CGCATCATCC	AAAGATCCTA	AGAAGAAAGA	2460
TAGCATAATA	TCTGTTAACT	CAATAGATAT	AGAAGTAACT	AAAGAGAGCC	CGATAACATC	2520
TGGCACCAAC	ATCATCAATC	CAACAAGTGA	AGCCGACAGT	ACCCCAAGAA	CAAAAGCCAA	2580
CTACCCAAGA	AAACCCCTAG	TAAGCTTCAA	AGAAGATCTC	ACCCCAAGTG	ACAACCCTTT	2640
TTCTAAGTTG	TACAAGGAAA	CAATAGAAAC	ATTTGATAAC	AATGAAGAAG	AATCTAGCTA	2700
CTCATATGAA	GAGATAAATG	ATCAAACAAA	TGACAACATT	ACAGCAAGAC	TAGATAGAAT	2760
TGATGAAAAA	TTAAGTGAAA	TATTAGGAAT	GCTCCATACA	TTAGTAGTTG	CAAGTGCAGG	2820
ACCCACTTCA	GCTCGCGATG	GAATAAGAGA	TGCTATGGTT	GGTCTAAGAG	AAGAGATGAT	2880
AGAAAAAATA	AGAGCGGAAG	CATTAATGAC	CAATGATAGG	TTAGAGGCTA	TGGCAAGACT	2940
TAGGAATGAG	GAAAGCGAAA	AAATGGCAAA	AGACACCTCA	GATGAAGTGT	CTCTTAATCC	3000
AACTTCCAAA	AAATTGAGTG	ACTTGTTGGA	AGACAACGAT	AGTGACAATG	ATCTATCACT	3060
TGATGATTTT	TGATCAGCGA	TCAACTCACT	CAGCAATCAA	CAACATCAAT	AAAACAGACA	3120
TCAATCCATT	GAATCAACTG	CCAGACCGAA	CAAACAAACG	TCCATCAGTA	GAACCACCAA	3180
CCAATCAATC	AACCAATTGA	TCAATCAGCA	ACCCGACAAA	ATTAACAATA	TAGTAACAAA	3240
AAAAGAACAA	GATGGGGCAA	ATATGGAAAC	ATACGTGAAC	AAGCTTCACG	AAGGCTCCAC	3300
ATACACAGCA	GCTGTTTCAGT	ACAATGTTCT	AGAAAAAGAT	GATGATCCTG	CATCACTAAC	3360

SUBSTITUTE SHEET (RULE 26)

- 349 -

AATATGGGTG CCTATGTTCC AGTCATCTGT GCCAGCAGAC TTGCTCATAA AAGAACTTGC	3420
AAGCATCAAT ATACTAGTGA AGCAGATCTC TACGCCCCAA GGACCTTCAC TACGAGTCAC	3480
GATTAACTCA AGAAGTGCTG TGCTGGCTCA AATGCCTAGT AATTTCATCA TAAGCGCAA	3540
TGTATCATTG GATGAAAGAA GCAAATTAGC ATATGATGTA ACTACACCTT GTGAAATCAA	3600
AGCATGCAGT CTAACATGCT TAAAAGTAAA AAGTATGTTA ACTACAGTCA AAGATCTTAC	3660
CATGAAGACA TTCAACCCCA CTCATGAGAT CATTGCTCTA TGTGAATTTG AAAATATTAT	3720
GACATCAAAA AGAGTAATAA TACCAACCTA TCTAAGATCA ATTAGTGTCA AGAACAAGGA	3780
TCTGAACTCA CTAGAAAATA TAGCAACCAC CGAATTCAAA AATGCTATCA CCAATGCAAA	3840
AATTATTCCT TAGTGCAGGAT TAGTGTTAGT TATCACAGTT ACTGACAATA AAGGAGCATT	3900
CAATATATC AAACACAGA GTCAATTTAT AGTAGATCTT GGTGCCTACC TAGAAAAAGA	3960
GAGCATATAT TATGTGACTA CTAATTGGAA GCATACAGCT ACACGTTTTT CAATCAAACC	4020
ACTAGAGGAT TAACTTAAT TATCAACACT GAATGACAGG TCCACATATA TCCTCAAAC	4080
ACACACTATA TCCAAACATC ATAAACATCT AACTACACA CTTATCACA CAAACCAATC	4140
CCACTCAAAA TCCAAAATCA CTACCAGCCA CTATCTGCTA GACCTAGAGT GCGAATAGGT	4200
AAATAAAACC AAAATATGGG GTAAATAGAC ATTAGTTAGA GTTCAATCAA TCTTAACAAC	4260
CATTTATACC GCCAATTCAA CACATATACT ATAAATCTTA AAATGGGAAA TACATCCATC	4320
ACAATAGAAT TCACAAGCAA ATTTTGGCCC TATTTTACAC TAATACATAT GATCTTAACT	4380
CTAATCTTTT TACTAATTAT AATCACTATT ATGATTGCAA TACTAAATAA GCTAAGTGAA	4440
CATAAAGCAT TCTGTAACAA AACTCTTGAA CTAGGACAGA TGTATCAAAT CAACACATAG	4500
AGTTCTACCA TTATGCTGTG TCAAATTATA ATCCTGTATA TATAAACAAA CAAATCCAAT	4560
CTTCTCACAG AGTCATGGTG TCGCAAAACC ACGCTAACTA TCATGGTAGC ATAGAGTAGT	4620
TATTTAAAAA TTAACATAAT GATGAATTGT TAGTATGAGA TCAAAAACAA CATTGGGGCA	4680
AATGCAACCA TGTCCAAACA CAAGAATCAA CGCACTGCCA GGACTCTAGA AAAGACCTGG	4740
GATACTCTTA ATCATCTAAT TGTAATATCC TCTTGTTTAT ACAGATTAAA TTTAAAATCT	4800
ATAGCACAAA TAGCACTATC AGTTTTGGCA ATGATAATCT CAACCTCTCT CATAATTGCA	4860
GCCATAATAT TCATCATCTC TGCCAATCAC AAAGTTACAC TAACAACGGT CACAGTTCAA	4920

SUBSTITUTE SHEET (RULE 26)

- 350 -

ACAATAAAAA ACCCACTGA AAAAAACATC ACCACCTACC CTACTCAAGT CTCACCAGAA	4980
AGGGTTAGTT CATCCAAGCA ACCCACAACC ACATCACCAA TCCACACAAG TTCAGCTACA	5040
ACATCACCCA ATACAAAATC AGAAACACAC CATACAACAG CACAAACCAA AGGCAGAACC	5100
ACCACTTCAA CACAGACCAA CAAGCCAAGC ACAAACCAC GTCCAAAAAA TCCACCAAAA	5160
AAAGATGATT ACCATTTTGA AGTGTTC AAC TTCGTTCCCT GCAGTATATG TGGCAACAAT	5220
CAACTTTGCA AATCCATCTG CAAAACAATA CCAAGCAACA AACCAGAGAA GAAACCAACC	5280
ATCAAACCCA CAAACAAACC AACCACCAAA ACCACAAACA AAAGAGACCC AAAAACACCA	5340
GCCAAACGA CGAAAAAGA AACTACCACC AACCCAACAA AAAAATAAC CCTCAAGACC	5400
ACAGAAAGAG ACACCAGCAC CTCACAATCC ACTGCACTCG ACACAACCAC ATTAAAAAC	5460
ACAGTCCAAC AGCAATCCCT CCTCTCAACC ACCCCCGAAA ACACACCCAA CTCCACACAA	5520
ACACCCACAG CATCCGAGCC CTCCACACCA AACTCCACCC AAAAACCCEA GCCACATGCT	5580
TAGTTATTCA AAAACTACAT CTTAGCAGAG AACCGTGATC TATCAAGCAA GAACGAAATT	5640
AAACCTGGGG CAAATAACCA TGGAGTTGAT GATCCACAAG TCAAGTGCAA TCTTCCTAAC	5700
TCTTGCTATT AATGCATTGT ACCTCACCTC AAGTCAGAAC ATAAGTGAGG AGTTTACCA	5760
ATCGACATGT AGTGCAGTTA GCAGAGGTTA TTTTAGTGCT TTAAGAACAG GTTGGTATAC	5820
TAGTGTGATA ACAATAGAAT TAAGTAATAT AAAAGAAACC AAATGCAATG GAACTGACAC	5880
TAAAGTAAAA CTTATGAAAC AAGAATTAGA TAAGTATAAG AATGCAGTAA CAGAATTACA	5940
GCTACTTATG CAAAACACAC CAGCTGTCAA CAACCGGGCC AGAAGAGAAG CACCACAGTA	6000
TATGAACTAC ACAATCAATA CCACTAAAAA CCTAAATGTA TCAATAAGCA AGAAGAGGAA	6060
ACGAAGATTT CTAGGCTTCT TGTTAGGTGT GGGATCTGCA ATAGCAAGTG GTATAGCTGT	6120
ATCAAAAGTT CTACACCTTG AAGGAGAAGT GAACAAGATC AAAAATGCTT TGTTGTCTAC	6180
AAACAAAGCT GTAGTCAGTT TATCAAATGG GGTCAAGTGT TTAACCAGCA AAGTGTTAGA	6240
TCTCAAGAAT TACATAAATA ACCAATTATT ACCCATAGTA AATCAACAGA GCTGTGCGAT	6300
CTCCAACATT GAAACAGTTA TAGAATTCCA GCAGAAGAAC AGCAGATTGT TGGAAATCAC	6360
CAGAGAATTT AGTGTCAATG CAGGTGTAAC AACACCTTTA AGCACTTACA TGTTGACAAA	6420
CAGTGAGTTA CTATCATTA TCAATGATAT GCCTATAACA AATGATCAGA AAAAATTAAT	6480

SUBSTITUTE SHEET (RULE 26)

- 351 -

GTCAAGCAAT GTTCAGATAG TAAGGCAACA AAGTTATTCC ATCATGTCTA TAATAAAGGA	6540
AGAAGTCCTT GCATATGTTG TACAGCTGCC TATCTATGGT GTAATAGATA CACCTTGCTG	6600
GAAATTGCAC ACATCGCCTC TATGCACTAC CAACATCAAA GAAGGATCAA ATATTTGTTT	6660
AACAAGGACT GATAGAGGAT GGTATTGTGA TAATGCAGGA TCAGTATCCT TCTTCCACA	6720
GGCTGACACT TGTAAAGTAC AGTCCAATCG AGTATTTTGT GACACTATGA ACAGTTTGAC	6780
ATTACCAAGT GAAGTCAGCC TTTGTAACAC TGACATATTC AATTCCAAGT ATGACTGCAA	6840
AATTATGACA TCAAAAACAG ACATAAGCAG CTCAGTAATT ACTTCTCTTG GAGCTATAGT	6900
GTCATGCTAT GGTAAACTA AATGCACTGC ATCCAACAAA AATCGTGGGA TTATAAAGAC	6960
ATTTTCTAAT GGTGTGACT ATGTGTCAA CAAAGGAGTA GATACTGTGT CAGTGGGCAA	7020
CACTTTATAC TATGTAAACA AGCTGGAAGG CAAGAACCTT TATGTAAAAG GGGAACCTAT	7080
AATAAATTAC TATGACCCTC TAGTGTTCCT TTCTGATGAG TTTGATGCAT CAATATCTCA	7140
AGTCAATGAA AAAATCAATC AAAGTTTAGC TTTTATTCGT AGATCTGATG AATTACTACA	7200
TAATGTAAAT ACTGGCAAAT CTACTACAAA TATTATGATA ACTACAATTA TTATAGTAAT	7260
CATTGTAGTA TTGTTATCAT TAATAGCTAT TGGTTTACTG TTGTATTGTA AAGCCAAAAA	7320
CACACCAGTT AACTAAGCA AAGACCAACT AAGTGGATC AATAATATTG CATTAGCAA	7380
ATAGACAAAA AACCACCTGA TCATGTTTCA ACAACAATCT GCTGACCACC AATCCCAAAT	7440
CAACTTACAA CAAATATTTT AACATCACAG TACAGGCTGA ATCATTTCCT CACATCATGC	7500
TACCCACATA ACTAAGCTAG ATCCTTAACT TATAGTTACA TAAAAACCTC AAGTATCACA	7560
ATCAACCACT AAATCAACAC ATCATTACACA AAATTAACAG CTGGGGCAAA TATGTCGCGA	7620
AGAAATCCTT GTAAATTTGA GATTAGAGGT CATTGCTTGA ATGGTAGAAG ATGCACTAC	7680
AGTCATAATT ACTTTGAATG GCCTCCTCAT GCATTACTAG TGAGGCAAAA CTTATGTGA	7740
AACAAGATAC TCAAGTCAAT GGACAAAAGC ATAGACACTT TGTCTGAAAT AAGTGGAGCT	7800
GCTGAACTGG ATAGAACAGA AGAATATGCT CTTGGTATAG TTGGAGTGCT AGAGAGTTAC	7860
ATAGGATCTA TAAACAACAT AACAAAACAA TCAGCATGTG TTGCTATGAG TAACTTCTT	7920
ATTGAGATCA ATAGTGATGA CATTAAAAAG CTTAGAGATA ATGAAGAACC CAATTCACCT	7980
AAGATAAGAG TGTACAATAC TGTTATATCA TACATTGAGA GCAATAGAAA AAACAACAAG	8040

SUBSTITUTE SHEET (RULE 26)

- 352 -

CAAACCATCC ATCTGCTCAA GAGACTACCA GCAGACGTGC TGAAGAAGAC AATAAAGAAC	8100
ACATTAGATA TCCACAAAAG CATAACCATA AGCAATCCAA AAGAGTCAAC TGTGAATGAT	8160
CAAAATGACC AAACCAAAAA TAATGATATT ACCGGATAAA TATCCTTGTA GTATATCATC	8220
CATATTGATC TCAAGTGAAA GCATGGTTGC TACATTCAAT CATAAAAACA TATTACAATT	8280
TAACCATAAC TATTTGGATA ACCACCAGCG TTTATTAAAT CATATATTTG ATGAAATTCA	8340
TTGGACACCT AAAAATTAT TAGATGCCAC TCAACAATTT CTCCAACATC TTAACATCCC	8400
TGAAGATATA TATACAGTAT ATATATTAGT GTCATAATGC TTGACCATAA CGACTCTATG	8460
TCATCCAACC ATAAAACTAT TTTGATAAGG TTATGGGACA AAATGGATCC CATTATTAAT	8520
GGAACTCTG CTAATGTGTA TCTAACTGAT AGTTATTTAA AAGGTGTTAT CTCTTTTCA	8580
GAGTGTAATG CTTTAGGGAG TTATCTTTTT AACGGCCCTT ATCTTAAAAA TGATTACACC	8640
AACTTAATTA GTAGACAAAG CCCACTACTA GAGCATATGA ATCTTAAAAA ACTAACTATA	8700
ACACAGTCAT TAATATCTAG ATATCATAAA GGTGAACTGA AATTAGAAGA ACCAACTTAT	8760
TTCCAGTCAT TACTTATGAC ATATAAAAGT ATGTCCTCGT CTGAACAAAT TGCTACAACT	8820
AACTTACTTA AAAAAATAAT ACGAAGAGCC ATAGAAATAA GTGATGTAAA GGTGTACGCC	8880
ATCTTGAATA AACTAGGATT AAAGGAAAAG GACAGAGTTA AGCCCAACAA TAATTCAGGT	8940
GATGAAAAC CAGTACTTAC AACTATAATT AAAGATGATA TACTTTCGGC TGTGGAAAAC	9000
AATCAATCAT ATACAAATC AGACAAAAGT CACTCAGTAA ATCAAAATAT CACTATCAAA	9060
ACAACACTCT TGAAAAAATT GATGTGTTCA ATGCAACATC CTCCATCATG GTTAATACAC	9120
TGGTTCAATT TATATACAAA ATTAAATAAC ATATTAACAC AATATCGATC AAATGAGGTA	9180
AAAAGTCATG GGTTTATATT AATAGATAAT CAACTTTTAA GTGGTTTTCA GTTTATTTTA	9240
AATCAATATG GTTGTATCGT TTATCATAAA GGA CTCAAAA AAATCACAAC TACTACTTAC	9300
AATCAATTTT TGACATGGAA AGACATCAGC CTTAGCAGAT TAAATGTTTG CTTAATTACT	9360
TGGATAAGTA ATTGTTTAAA TACATTAAAC AAAAGCTTAG GGCTGAGATG TGGATTCAAT	9420
AATGTTGTGT TATCACAATT ATTTCTTTAT GGAGATTGTA TACTGAAATT ATTTCATAA	9480
GAAGGCTTCT ACATAATAAA AGAAGTAGAG GGATTTATTA TGTCTTTAAT TCTAAACATA	9540
ACAGAAGAAG ATCAATTTAG GAAACGATTT TATAATAGCA TGCTAAATAA CATCACAGAT	9600

SUBSTITUTE SHEET (RULE 26)



- 353 -

GCAGCTATTA	AGGCTCAAAA	GGACCTACTA	TCAAGAGTAT	GTCACACTTT	ATTAGACAAG	9660
ACAGTGTCTG	ATAATATCAT	AAATGGTAAA	TGGATAATCC	TATTAAGTAA	ATTTCTTAAA	9720
TTGATTAAGC	TTGCAGGTGA	TAATAATCTC	AATAACTTGA	GTGAGCTATA	TTTTCTCTTC	9780
AGAATCTTTG	GACATCCAAT	GGTCGATGAA	AGACAAGCAA	TGGATTCTGT	AAGAATTAAC	9840
TGTAATGAAA	CTAAGTTCTA	CTTATTAAGT	AGTCTAAGTA	CATTAAGAGG	TGCTTTCATT	9900
TATAGAATCA	TAAAAGGGTT	TGTAAATACC	TACAACAGAT	GGCCACCTT	AAGGAATGCT	9960
ATTGTCTTAC	CTCTAAGATG	GTTAACTAC	TATAAACTTA	ATACTTATCC	ATCTCTACTT	10020
GAAATCACAG	AAAATGATTT	GATTATTTTA	TCAGGATTGC	GGTTCTATCG	TGAGTTTCAT	10080
CTGCCTAAAA	AAGTGGATCT	TGAAATGATA	ATAAATGACA	AAGCCATTTC	ACCTCCAAAA	10140
GATCTAATAT	GGACTAGTTT	TCCTAGAAAT	TACATGCCAT	CACATATACA	AAATTATATA	10200
GAACATGAAA	AGTTGAAGTT	CTCTGAAAGC	GACAGATCGA	GAAGAGTACT	AGAGTATTAC	10260
TTGAGAGATA	ATAAATTCAA	TGAATGCGAT	CTATACAATT	GTGTAGTCAA	TCAAAGCTAT	10320
CTCAACAAC	CTAATCACGT	GGTATCACTA	ACTGGTAAAG	AAAGAGAGCT	CAGTGTAGGT	10380
AGAATGTTTG	CTATGCAACC	AGGTATGTTT	AGGCAAATCC	AAATCTTAGC	AGAGAAAATG	10440
ATAGCTGAAA	ATATTTTACA	ATTCTTCCCT	GAGAGTTTGA	CAAGATATGG	TGATCTAGAG	10500
CTTCAAAAGA	TATTAGAATT	AAAAGCAGGA	ATAAGCAACA	AGTCAAATCG	TTATAATGAT	10560
AACTACAACA	ATTATATCAG	TAAATGTTCT	ATCATTACAG	ATCTTAGCAA	ATTCAATCAG	10620
GCATTTAGAT	ATGAAACATC	ATGTATCTGC	AGTGATGTAT	TAGATGAACT	GCATGGAGTA	10680
CAATCTCTGT	TCTCTTGGTT	GCATTTAACA	ATACCTCTTG	TCACAATAAT	ATGTACATAT	10740
AGACATGCAC	CTCCTTTCAT	AAAGGATCAT	GTTGTTAATC	TTAATGAGGT	TGATGAACAA	10800
AGTGGATTAT	ACAGATATCA	TATGGGTGGT	ATTGAGGGCT	GGTGTCAAAA	ACTGTGGACC	10860
ATTGAAGCTA	TATCATTATT	AGATCTAATA	TCTCTCAAAG	GGAAATTCTC	TATCACAGCT	10920
CTGATAAATG	GTGATAATCA	GTCAATTGAT	ATAAGCAAAC	CAGTTAGACT	TATAGAGGGT	10980
CAGACCCATG	CACAAGCAGA	TTATTTGTTA	GCATTAAATA	GCCTTAAATT	GTTATATAAA	11040
GAGTATGCAG	GTATAGGCCA	TAAGCTTAAG	GGAACAGAGA	CCTATATATC	CCGAGATATG	11100
CAGTTCATGA	GCAAAACAAT	CCAGCACAAAT	GGAGTGTACT	ATCCAGCCAG	TATCAAAAAA	11160

SUBSTITUTE SHEET (RULE 26)

- 354 -

GTCCTGAGAG TAGGTCCATG GATAAACACG ATACTTGATG ATTTTAAAGT TAGTTTAGAA 11220  
TCTATAGGCA GCTTAACACA GGAGTTAGAA TACAGAGGAG AAAGCTTATT ATGCAGTTTA 11280  
ATATTTAGGA ACATTTGGTT ATACAATCAA ATTGCTTTGC AACTCCGAAA TCATGCATTA 11340  
TGTAACAATA AGCTATATTT AGATATATTG AAAGTATTAA AACACTTAAA AACTTTTTTTT 11400  
AATCTTGATA GCATTGATAT GGCTTTATCA TTGTATATGA ATTTGCCTAT GCTGTTTGGT 11460  
GGTGGTGATC CTAATTTGTT ATATCGAAGC TTTTATAGGA GAACTCCAGA CTTCCCTTACA 11520  
GAAGCTATAG TACATTCAGT GTTTGTGTTG AGCTATTATA CTGGTCACGA TTTACAAGAT 11580  
AAGCTCCAGG ATCTTCCAGA TGATAGACTG AACAAATTCT TGACATGTGT CATCACATTT 11640  
GATAAAAATC CCAATGCCGA GTTTGTAACA TTGATGAGGG ATCCACAGGC TTTAGGGTCT 11700  
GAAAGGCAAG CTAAAATTAC TAGTGAGATT AATAGATTAG CAGTAACAGA AGTCTTAAGT 11760  
ATAGCCCCAA ACAAATATT TTCTAAAAGT GCACAACATT ATACTACCAC TGAGATTGAT 11820  
CTAAATGACA TTATGCAAAA TATAGAACCA ACTTACCCTC ATGGATTAAG AGTTGTTTAT 11880  
GAAAGTTTAC CTTTTTATAA AGCAGAAAAA ATAGTTAATC TTATATCAGG AACAAAATCC 11940  
ATAACTAATA TACTTGAAAA AACATCAGCA ATAGATACAA CTGATATTAA TAGGGCTACT 12000  
GATATGATGA GGAAAAATAT AACTTTACTT ATAAGGATAC TTCCACTAGA TTGTAACAAA 12060  
GACAAAAGAG AGTTATTAAG TTTAGAAAAT CTTAGTATAA CTGAATTAAG CAAGTATGTA 12120  
AGAGAAAGAT CTTGTCATT ATCCAATATA GTAGGAGTAA CATCGCCAAG TATTATGTTT 12180  
ACAATGGACA TTAAATATAC AACTAGCACT ATAGCCAGTG GTATAATAAT AGAAAAATAT 12240  
AATGTTAATA GTTTAACTCG TGGTGAAAGA GGACCCACCA AGCCATGGGT AGGCTCATCC 12300  
ACGCAGGAGA AAAAAACAAT GCCAGTGTAC AACAGACAAG TTTTAACCAA AAAGCAAAGA 12360  
GACCAAATAG ATTTATTAGC AAAATTAGAC TGGGTATATG CATCCATAGA CAACAAAGAT 12420  
GAATTCATGG AAGAACTGAG TACTGGAACA CTTGGACTGT CATATGAAAA AGCCAAAAAG 12480  
TTGTTTCCAC AATATCTAAG TGTCAATTAT TTACACCGTT TAACAGTCAG TAGTAGACCA 12540  
TGTGAATTCC CTGCATCAAT ACCAGCTTAT AGAACAACAA ATTATCATTT TGATACTAGT 12600  
CCTATCAATC ATGTATTAAC AGAAAAGTAT GGAGATGAAG ATATCGACAT TGTGTTTCAA 12660  
AATTGCATAA GTTTTGGTCT TAGCCTGATG TCGGTTGTGG AACAAATTCAC AACATATGT 12720

SUBSTITUTE SHEET (RULE 26)

- 355 -

CCTAATAGAA TTATTCTCAT ACCGAAGCTG AATGAGATAC ATTTGATGAA ACCTCCTATA 12780  
TTTACAGGAG ATGTTGATAT CATCAAGTTG AAGCAAGTGA TACAAAAGCA GCACATGTTT 12840  
CTACCAGATA AAATAAGTTT AACCCAATAT GTAGAATTAT TCTTAAGTAA CAAAGCACTT 12900  
AAATCTGGAT CTCACATCAA CTCTAATTTA ATATTAGTAC ATAAAATGTC TGATTATTTT 12960  
CATAATGCTT ATATTTTAAG TACTAATTTA GCTGGACATT GGATTCTGAT TATTCAACTT 13020  
ATGAAAGATT CAAAAGGTAT TTTTGAAAAA GATTGGGGAG AGGGGTACAT AACTGATCAT 13080  
ATGTTCATTA ATTTGAATGT TTTCTTTAAT GCTTATAAGA CTTATTTGCT ATGTTTTTAT 13140  
AAAGGTTATG GTAAAGCAAA ATTAGAATGT GATATGAACA CTTCAGATCT TCTTTGTGTT 13200  
TTGGAGTTAA TAGACAGTAG CTACTGGAAA TCTATGTCTA AAGTTTTTCT AGAACAAAAA 13260  
GTCATAAAAT ACATAGTCAA TCAAGACACA AGTTTGCCTA GAATAAAAGG CTGTCACAGT 13320  
TTTAAGTTGT GGTTTTAAA ACGCCTTAAT AATGCTAAAT TTACCGTATG CCCTTGGGTT 13380  
GTTAACATAG ATTATCACCC AACACACATG AAAGCTATAT TATCTTACAT AGATTTAGTT 13440  
AGAAATGGGGT TAATAAATGT AGATAAATTA ACCATTAAAA ATAAAAACAA ATTCAATGAT 13500  
GAATTTTACA CATCAAATCT CTTTACATT AGTTATAACT TTTCAGACAA CACTCATTTG 13560  
CTAACAAAAC AAATAAGAAT TGCTAATTCA GAATTAGAAG ATAATTATAA CAAACTATAT 13620  
CACCCAACCC CAGAAACTTT AGAAAATATG TCATTAATTC CTGTTAAAAG TAATAATAGT 13680  
AACAAACCTA AATTTTGTAT AAGTGGAAAT ACCGAATCTA TGATGATGTC AACATTCTCT 13740  
AGTAAAATGC ATATTAAATC TTCCACTGTT ACCACAAGAT TCAATTATAG CAAACAAGAC 13800  
TTGTACAATT TATTTCCAAT TGTTGTGATA GACAAGATTA TAGATCATTC AGGTAATACA 13860  
GCAAAATCTA ACCAACTTTA CACCACCACT TCACATCAGA CATCTTTAGT AAGGAATAGT 13920  
GCATCACTTT ATTGCATGCT TCCTTGGCAT CATGTCAATA GATTTAACTT TGTATTTAGT 13980  
TCCACAGGAT GCAAGATCAG TATAGAGTAT ATTTTAAAAG ATCTTAAGAT TAAGGACCCC 14040  
AGTTGTATAG CATTATAGG TGAAGGAGCT GGTAACCTAT TATTACGTAC GGTAGTAGAA 14100  
CTTCATCCAG ACATAAGATA CATTTACAGA AGTTTAAAAG ATTGCAATGA TCATAGTTTA 14160  
CCTATTGAAT TTCTAAGGTT ATACAACGGG CATATAACA TAGATTATGG TGAGAATTTA 14220  
ACCATTCTCTG CTACAGATGC AACTAATAAC ATTCATTGGT CTTATTTACA TATAAAATTT 14280

SUBSTITUTE SHEET (RULE 26)

- 356 -

GCAGAACCTA TTAGCATCTT TGTCTGCGAT GCTGAATTAC CTGTTACAGC CAATTGGAGT 14340  
 AAAATTATAA TTGAATGGAG TAAGCATGTA AGAAAGTGCA AGTACTGTTC TTCTGTAAAT 14400  
 AGATGCATTT TAATTGCAAA ATATCATGCT CAAGATGACA TTGATTTCAA ATTAGATAAC 14460  
 ATTACTATAT TAAAACTTA CGTGTGCCTA GGTAGCAAGT TAAAAGGATC TGAAGTTTAC 14520  
 TTAATCCTTA CAATAGGCCC TGCAAATATA CTTCTGTGTT TTGATGTTGT ACAAAATGCT 14580  
 AAATTGATAC TTTCAAGAAC TAAAAATTC ATTATGCCTA AAAAACTGA CAAGGAATCT 14640  
 ATCGATGCAG ATATTAAAAG CTTAATACCT TTCCTTTGTT ACCCTATAAC AAAAAAGGA 14700  
 ATTAAGACTT CATTGTCAAA ATTGAAGAGT GTAGTTAATG GAGATATATT ATCATATTCT 14760  
 ATAGCTGGAC GTAATGAAGT ATTCAGCAAC AAGCTTATAA ACCACAAGCA TATGAATATC 14820  
 CTAAATGGC TAGATCATGT TTTAAATTTT AGATCAGCTG AACTTAATTA CAATCATTTA 14880  
 TACATGATAG AGTCCACATA TCCTTACTTA AGTGAATTGT TAAATAGTTT AACAACCAAT 14940  
 GAGCTCAAGA AGCTGATTAA AATAACAGGT AGTGTGCTAT ACAACCTTCC CAACGAACAG 15000  
 TAGTTTAAAA TATCATTAAAC AAGTTTGGTC AAATTTAGAT GCTAACACAT CATTATATTA 15060  
 TAGTTATTAA AAAATATACA AACTTTTCAA TAATTTAGCA TATTGATTCC AAAATTATCA 15120  
 TTTTAGTCTT AAGGGGTAA ATAAAAGTCT AAAACTAACA ATTATACATG TGCATTCACA 15180  
 ACACAACGAG ACATTAGTTT TTGACACTTT TTTTCTCGT 15219

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp  
 1 5 10 15

Ser Tyr Leu Lys Gly Val Ile Ser Phe Ser Glu Cys Asn Ala Leu Gly

SUBSTITUTE SHEET (RULE 26)

- 357 -

20					25					30						
Ser	Tyr	Leu	Phe	Asn	Gly	Pro	Tyr	Leu	Lys	Asn	Asp	Tyr	Thr	Asn	Leu	
35					40					45						
Ile	Ser	Arg	Gln	Ser	Pro	Leu	Leu	Glu	His	Met	Asn	Leu	Lys	Lys	Leu	
50					55					60						
Thr	Ile	Thr	Gln	Ser	Leu	Ile	Ser	Arg	Tyr	His	Lys	Gly	Glu	Leu	Lys	
65					70					75					80	
Leu	Glu	Glu	Pro	Thr	Tyr	Phe	Gln	Ser	Leu	Leu	Met	Thr	Tyr	Lys	Ser	
85					90					95						
Met	Ser	Ser	Ser	Glu	Gln	Ile	Ala	Thr	Thr	Asn	Leu	Leu	Lys	Lys	Ile	
100					105					110						
Ile	Arg	Arg	Ala	Ile	Glu	Ile	Ser	Asp	Val	Lys	Val	Tyr	Ala	Ile	Leu	
115					120					125						
Asn	Lys	Leu	Gly	Leu	Lys	Glu	Lys	Asp	Arg	Val	Lys	Pro	Asn	Asn	Asn	
130					135					140						
Ser	Gly	Asp	Glu	Asn	Ser	Val	Leu	Thr	Thr	Ile	Ile	Lys	Asp	Asp	Ile	
145					150					155					160	
Leu	Ser	Ala	Val	Glu	Asn	Asn	Gln	Ser	Tyr	Thr	Asn	Ser	Asp	Lys	Ser	
165					170					175						
His	Ser	Val	Asn	Gln	Asn	Ile	Thr	Ile	Lys	Thr	Thr	Leu	Leu	Lys	Lys	
180					185					190						
Leu	Met	Cys	Ser	Met	Gln	His	Pro	Pro	Ser	Trp	Leu	Ile	His	Trp	Phe	
195					200					205						
Asn	Leu	Tyr	Thr	Lys	Leu	Asn	Asn	Ile	Leu	Thr	Gln	Tyr	Arg	Ser	Asn	
210					215					220						
Glu	Val	Lys	Ser	His	Gly	Phe	Ile	Leu	Ile	Asp	Asn	Gln	Thr	Leu	Ser	
225					230					235					240	
Gly	Phe	Gln	Phe	Ile	Leu	Asn	Gln	Tyr	Gly	Cys	Ile	Val	Tyr	His	Lys	
245					250					255						
Gly	Leu	Lys	Lys	Ile	Thr	Thr	Thr	Thr	Tyr	Asn	Gln	Phe	Leu	Thr	Trp	
260					265					270						
Lys	Asp	Ile	Ser	Leu	Ser	Arg	Leu	Asn	Val	Cys	Leu	Ile	Thr	Trp	Ile	
275					280					285						
Ser	Asn	Cys	Leu	Asn	Thr	Leu	Asn	Lys	Ser	Leu	Gly	Leu	Arg	Cys	Gly	
290					295					300						

SUBSTITUTE SHEET (RULE 26)

- 358 -

Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile  
 305 310 315 320  
 Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu  
 325 330 335  
 Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe  
 340 345 350  
 Arg Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala  
 355 360 365  
 Ile Lys Ala Gln Lys Asp Leu Leu Ser Arg Val Cys His Thr Leu Leu  
 370 375 380  
 Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu  
 385 390 395 400  
 Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu  
 405 410 415  
 Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro  
 420 425 430  
 Met Val Asp Glu Arg Gln Ala Met Asp Ser Val Arg Ile Asn Cys Asn  
 435 440 445  
 Glu Thr Lys Phe Tyr Leu Leu Ser Ser Leu Ser Thr Leu Arg Gly Ala  
 450 455 460  
 Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp  
 465 470 475 480  
 Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr  
 485 490 495  
 Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Asn Asp  
 500 505 510  
 Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro  
 515 520 525  
 Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro  
 530 535 540  
 Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser  
 545 550 555 560  
 His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser  
 565 570 575

SUBSTITUTE SHEET (RULE 26)

- 359 -

Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe  
 580 585 590  
 Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn  
 595 600 605  
 Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser  
 610 615 620  
 Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln  
 625 630 635 640  
 Ile Leu Ala Glu Lys Met Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro  
 645 650 655  
 Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu  
 660 665 670  
 Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr  
 675 680 685  
 Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe  
 690 695 700  
 Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu  
 705 710 715 720  
 Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr  
 725 730 735  
 Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe  
 740 745 750  
 Ile Lys Asp His Val Val Asn Leu Asn Glu Val Asp Glu Gln Ser Gly  
 755 760 765  
 Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu  
 770 775 780  
 Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly  
 785 790 795 800  
 Lys Phe Ser Ile Thr Ala Leu Ile Asn Gly Asp Asn Gln Ser Ile Asp  
 805 810 815  
 Ile Ser Lys Pro Val Arg Leu Ile Glu Gly Gln Thr His Ala Gln Ala  
 820 825 830  
 Asp Tyr Leu Leu Ala Leu Asn Ser Leu Lys Leu Leu Tyr Lys Glu Tyr  
 835 840 845  
 Ala Gly Ile Gly His Lys Leu Lys Gly Thr Glu Thr Tyr Ile Ser Arg

SUBSTITUTE SHEET (RULE 26)

- 360 -

850	855	860
Asp Met Gln Phe Met Ser Lys Thr Ile Gln His Asn Gly Val Tyr Tyr		
865	870	875 880
Pro Ala Ser Ile Lys Lys Val Leu Arg Val Gly Pro Trp Ile Asn Thr		
885	890	895
Ile Leu Asp Asp Phe Lys Val Ser Leu Glu Ser Ile Gly Ser Leu Thr		
900	905	910
Gln Glu Leu Glu Tyr Arg Gly Glu Ser Leu Leu Cys Ser Leu Ile Phe		
915	920	925
Arg Asn Ile Trp Leu Tyr Asn Gln Ile Ala Leu Gln Leu Arg Asn His		
930	935	940
Ala Leu Cys Asn Asn Lys Leu Tyr Leu Asp Ile Leu Lys Val Leu Lys		
945	950	955 960
His Leu Lys Thr Phe Phe Asn Leu Asp Ser Ile Asp Met Ala Leu Ser		
965	970	975
Leu Tyr Met Asn Leu Pro Met Leu Phe Gly Gly Gly Asp Pro Asn Leu		
980	985	990
Leu Tyr Arg Ser Phe Tyr Arg Arg Thr Pro Asp Phe Leu Thr Glu Ala		
995	1000	1005
Ile Val His Ser Val Phe Val Leu Ser Tyr Tyr Thr Gly His Asp Leu		
1010	1015	1020
Gln Asp Lys Leu Gln Asp Leu Pro Asp Asp Arg Leu Asn Lys Phe Leu		
1025	1030	1035 1040
Thr Cys Val Ile Thr Phe Asp Lys Asn Pro Asn Ala Glu Phe Val Thr		
1045	1050	1055
Leu Met Arg Asp Pro Gln Ala Leu Gly Ser Glu Arg Gln Ala Lys Ile		
1060	1065	1070
Thr Ser Glu Ile Asn Arg Leu Ala Val Thr Glu Val Leu Ser Ile Ala		
1075	1080	1085
Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu		
1090	1095	1100
Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His		
1105	1110	1115 1120
Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys		
1125	1130	1135

SUBSTITUTE SHEET (RULE 26)



- 361 -

Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu  
 1140 1145 1150

Lys Thr Ser Ala Ile Asp Thr Thr Asp Ile Asn Arg Ala Thr Asp Met  
 1155 1160 1165

Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys  
 1170 1175 1180

Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr  
 1185 1190 1195 1200

Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile  
 1205 1210 1215

Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asp Ile Lys Tyr  
 1220 1225 1230

Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val  
 1235 1240 1245

Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly  
 1250 1255 1260

Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val  
 1265 1270 1275 1280

Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp  
 1285 1290 1295

Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu  
 1300 1305 1310

Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe  
 1315 1320 1325

Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser  
 1330 1335 1340

Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn  
 1345 1350 1355 1360

Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr  
 1365 1370 1375

Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly  
 1380 1385 1390

Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn  
 1395 1400 1405

SUBSTITUTE SHEET (RULE 26)

- 362 -

Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro  
 1410 1415 1420  
 Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile  
 1425 1430 1435 1440  
 Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr  
 1445 1450 1455  
 Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile  
 1460 1465 1470  
 Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn  
 1475 1480 1485  
 Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile  
 1490 1495 1500  
 Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu  
 1505 1510 1515 1520  
 Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn  
 1525 1530 1535  
 Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala  
 1540 1545 1550  
 Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu  
 1555 1560 1565  
 Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu  
 1570 1575 1580  
 Gln Lys Val Ile Lys Tyr Ile Val Asn Gln Asp Thr Ser Leu Arg Arg  
 1585 1590 1595 1600  
 Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asn  
 1605 1610 1615  
 Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His  
 1620 1625 1630  
 Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met  
 1635 1640 1645  
 Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe  
 1650 1655 1660  
 Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe  
 1665 1670 1675 1680  
 Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser

SUBSTITUTE SHEET (RULE 26)

- 363 -

1685	1690	1695
Glu Leu Glu Asp Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr 1700	1705	1710
Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys 1715	1720	1725
Pro Lys Phe Cys Ile Ser Gly Asn Thr Glu Ser Met Met Met Ser Thr 1730	1735	1740
Phe Ser Ser Lys Met His Ile Lys Ser Ser Thr Val Thr Thr Arg Phe 1745	1750	1755 1760
Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile 1765	1770	1775
Asp Lys Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu 1780	1785	1790
Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser 1795	1800	1805
Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val 1810	1815	1820
Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp 1825	1830	1835 1840
Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala 1845	1850	1855
Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg 1860	1865	1870
Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile 1875	1880	1885
Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu 1890	1895	1900
Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser 1905	1910	1915 1920
Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp 1925	1930	1935
Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp 1940	1945	1950
Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys 1955	1960	1965

SUBSTITUTE SHEET (RULE 26)

- 364 -

Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu  
 1970 1975 1980  
 Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu  
 1985 1990 1995 2000  
 Lys Gly Ser Glu Val Tyr Leu Ile Leu Thr Ile Gly Pro Ala Asn Ile  
 2005 2010 2015  
 Leu Pro Val Phe Asp Val Val Gln Asn Ala Lys Leu Ile Leu Ser Arg  
 2020 2025 2030  
 Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp  
 2035 2040 2045  
 Ala Asp Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys  
 2050 2055 2060  
 Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Asn Gly  
 2065 2070 2075 2080  
 Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn  
 2085 2090 2095  
 Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His  
 2100 2105 2110  
 Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met  
 2115 2120 2125  
 Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr  
 2130 2135 2140  
 Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr  
 2145 2150 2155 2160  
 Asn Leu Pro Asn Glu Gln  
 2165

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

SUBSTITUTE SHEET (RULE 26)

- 365 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ACGGGAAAAA AATGCGTACT ACAAACCTTGC ACATTCGAAA AAAATGGGGC AAATAAGAAC	60
TTGATAAGTG CTATTTAAGT CTAACCTTTT CAATCAGAAA TGGGGTGCAA TTCACTGAGC	120
ATGATAAAGG TTAGATTACA AAATTTATTT GACAATGACG AAGTAGCATT GTTAAAAATA	180
ACATGTTATA CTGATAAATT AATTCTTCTG ACCAATGCAT TAGCCAAAGC AGCAATACAT	240
ACAATTAAAT TAAACGGCAT AGTTTTTATA CATGTTATAA CAAGCAGTGA AGTGTGCCCT	300
GATAACAATA TTGTAGTGAA ATCTAACTTT ACAACAATGC CAATACTACA AAATGGAGGA	360
TACATATGGG AATTGATTGA GTTGACACAC TGCTCTCAAT TAAACGGTTT AATGGATGAT	420
AATTGTGAAA TCAAATTTTC TAAAAGACTA AGTGACTCAG TAATGACTAA TTATATGAAT	480
CAAATATCTG ACTTACTTGG GCTTGATCTC AATTCATGAA TTATGTTTAG TCTAATTCAA	540
TAGACATGTG TTTATTACCA TTTTAGTTAA TATAAAAACT CATCAAAGGG AAATGGGGCA	600
AATAAACTCA CCTAATCAAT CAAACCATGA GCACTACAAA TGACAACACT ACTATGCAAA	660
GATTGATGAT CACAGACATG AGACCCCTGT CAATGGATTG AATAATAACA TCTCTTACCA	720
AAGAAATCAT CACACACAAA TTCATATACT TGATAAACAA TGAATGTATT GTAAGAAAAC	780
TTGATGAAAG ACAAGCTACA TTTACATTCT TAGTCAATTA TGAGATGAAG CTAATGCACA	840
AAGTAGGGAG TACCAAATAC AAAAAATACA CTGAATATAA TACAAAATAT GGCACCTTCC	900
CCATGCCTAT ATTTATCAAT CACGGCGGGT TTCTAGAATG TATTGGCATT AAGCCTACAA	960
AACACACTCC TATAATATAC AAATATGACC TCAACCCGTG AATTCCAACA AAAAAACCAA	1020
CCCAACCAAA CCAAATATT CCTCAAACAA CAGTGCTCAA TAGTTAAGAA GGAGCTAATC	1080
CATTTTAGTA ATTA AAAAATA AAAGTAAAGC CAATAACATA AATTGGGGCA AATACAAAGA	1140
TGGCTCTTAG CAAAGTCAAG TTGAATGATA CATTAAATAA GGATCAGCTG CTGTCATCCA	1200
GCAAATACAC TATTCAACGT AGTACAGGAG ATAATATTGA CACTCCCAAT TATGATGTGC	1260
AAAAACACCT AAACAAACTA TGTGGTATGC TATTAATCAC TGAAGATGCA AATCATAAAT	1320
TCACAGGATT AATAGGTATG TTATATGCTA TGTCCAGGTT AGGAAGGGAA GACACTATAA	1380
AGATACTTAA AGATGCTGGA TATCATGTTA AAGCTAATGG AGTAGATATA ACAACATATC	1440
GTCAAGATAT AAATGGAAAG GAAATGAAAT TCGAAGTATT AACATTATCA AGCTTGACAT	1500

SUBSTITUTE SHEET (RULE 26)

- 366 -

CAGAAATACA AGTCAATATT GAGATAGAAT CTAGAAAGTC CTACAAAAAA ATGCTAAAAG	1560
AGATGGGAGA AGTGGCTCCA GAATATAGGC ATGATTCTCC AGACTGTGGG ATGATAATAC	1620
TGTGTATAGC TGCACCTTGTG ATAACCAAT TAGCAGCAGG AGACAGATCA GGTCTTACAG	1680
CAGTAATTAG GAGGGCAAAC AATGTCTTAA AAAACGAAAT AAAACGATAC AAGGGCCTCA	1740
TACCAAAGGA TATAGCTAAC AGTTTTTATG AAGTGTGTTGA AAAACACCCT CATCTTATAG	1800
ATGTTTTCGT GCACTTTGGC ATTGCACAAT CATCCACAAG AGGGGGTAGT AGAGTTGAAG	1860
GAATCTTTGC AGGATTGTTT ATGAATGCCT ATGGTTCAGG GCAAGTAATG CTAAGATGGG	1920
GAGTTTTAGC CAAATCTGTA AAAAATATCA TGCTAGGACA TGCTAGTGTC CAGGCAGAAA	1980
TGGAGCAAGT TGTGGAAGTC TATGAGTATG CACAGAAGTT GGGAGGAGAA GCTGGATTCT	2040
ACCATATATT GAACAATCCA AAAGCATCAT TGCTGTCATT AACTCAATTT CCCAACTTCT	2100
CAAGTGTGGT CCTAGGCAAT GCAGCAGGTC TAGGCATAAT GGGAGAGTAT AGAGGTACAC	2160
CAAGAAACCA GGATCTTTAT GATGCAGCTA AAGCATATGC AGAGCAACTC AAAGAAAATG	2220
GAGTAATAAA CTACAGTGTA TTAGACTTAA CAGCAGAAGA ATTGGAAGCC ATAAAGCATC	2280
AACTCAACCC CAAAGAAGAT GATGTAGAGC TTTAAGTTAA CAAAAAATAC GGGGCAAATA	2340
AGTCAACATG GAGAAGTTTG CACCTGAATT TCATGGAGAA GATGCAAATA ACAAGCTAC	2400
CAAATTCCTA GAATCAATAA AGGGCAAGTT CGCATCATCC AAAGATCCTA AGAAGAAAGA	2460
TAGCATAATA TCTGTAACT CAATAGATAT AGAAGTAACT AAAGAGAGCC CGATAACATC	2520
TGGCACCAAC ATCATCAATC CAACAAGTGA AGCCGACAGT ACCCCAGAAA CAAAAGCCAA	2580
CTACCCAAGA AAACCCCTAG TAAGCTTCAA AGAAGATCTC ACCCCAAGTG ACAACCCCTT	2640
TTCTAAGTTG TACAAGGAAA CAATAGAAAC ATTTGATAAC AATGAAGAAG AATCTAGCTA	2700
CTCATATGAA GAGATAAATG ATCAAACAAA TGACAACATT ACAGCAAGAC TAGATAGAA	2760
TGATGAAAAA TTAAGTGAAA TATTAGGAAT GCTCCATACA TTAGTAGTTG CAAGTGCAGG	2820
ACCCACTTCA GCTCGCGATG GAATAAGAGA TGCTATGGTT GGTCTAAGAG AAGAGATGAT	2880
AGAAAAATA AGAGCGGAAG CATTAAATGAC CAATGATAGG TTAGAGGCTA TGGCAAGACT	2940
TAGGAATGAG GAAAGCGAAA AAATGGCAA AGACACCTCA GATGAAGTGT CTCTTAATCC	3000
AACTTCCAAA AAATTGAGTG ACTTGTTGGA AGACAACGAT AGTGACAATG ATCTATCACT	3060

SUBSTITUTE SHEET (RULE 26)

- 367 -

TGATGATTTT TGATCAGCGA TCAACTCACT CAGCAATCAA CAACATCAAT AAAACAGACA 3120  
TCAATCCATT GAATCAACTG CCAGACCGAA CAAACAAACG TCCATCAGTA GAACCACCAA 3180  
CCAATCAATC AACCAATTGA TCAATCAGCA ACCCGACAAA ATTAACAATA TAGTAACAAA 3240  
AAAAGAACAA GATGGGGCAA ATATGGAAAC ATACGTGAAC AAGCTTCACG AAGGCTCCAC 3300  
ATACACAGCA GCTGTTTCAGT ACAATGTTCT AGAAAAAGAT GATGATCCTG CATCACTAAC 3360  
AATATGGGTG CCTATGTTCC AGTCATCTGT GCCAGCAGAC TTGCTCATAA AAGAACTTGC 3420  
AAGCATCAAT ATACTAGTGA AGCAGATCTC TACGCCCAAA GGACCTTCAC TACGAGTCAC 3480  
GATTAACTCA AGAAGTGCTG TGCTGGCTCA AATGCCTAGT AATTTTCATCA TAAGCGCAAA 3540  
TGTATCATTA GATGAAAGAA GCAAATTAGC ATATGATGTA ACTACACCTT GTGAAATCAA 3600  
AGCATGCAGT CTAACATGCT TAAAAGTAAA AAGTATGTTA ACTACAGTCA AAGATCTTAC 3660  
CATGAAGACA TTCAACCCCA CTCATGAGAT CATTGCTCTA TGTGAATTTG AAAATATTAT 3720  
GACATCAAAA AGAGTAATAA TACCAACCTA TCTAAGATCA ATTAGTGTCA AGAACAAGGA 3780  
TCTGAACTCA CTAGAAAATA TAGCAACCAC CGAATTCAAA AATGCTATCA CCAATGCAAA 3840  
AATTATTCCT TATGCAGGAT TAGTGTTAGT TATCACAGTT ACTGACAATA AAGGAGCATT 3900  
CAAATATATC AAACCACAGA GTCAATTTAT AGTAGATCTT GGTGCCTACC TAGAAAAAGA 3960  
GAGCATATAT TATGTGACTA CTAATTGGAA GCATACAGCT ACACGTTTTT CAATCAAACC 4020  
ACTAGAGGAT TAACTTAAT TATCAACACT GAATGACAGG TCCACATATA TCCTCAAAC 4080  
ACACACTATA TCCAAACATC ATAAACATCT ACACTACACA CTTTCATCACA CAAACCAATC 4140  
CCACTCAAAA TCCAAAATCA CTACCAGCCA CTATCCGCTA GACCTAGAGT GCGAATAGGC 4200  
AAATAAAACC AAAATATGGG GTAAATAGAC ATTAGTTAGA GTTCAATCAA TCTTAACAAC 4260  
CATTTATACC GCCAATTCAA CACATATACT ATAAATCTTA AAATGGGAAA TACATCCATC 4320  
ACAATAGAAC TCACAAGCAA ATTTTGGCCC TATTTTACAC TAATACATAT GATCTTAACT 4380  
CTAATCTTTT TACTAATTAT AATCACTATC ATGATTGCAA CACTAAATAA GCTAAGTGAA 4440  
CACAAAGCAT TCTGCAACAA AACTCTTGAA CTAGGACAGA TGTACCAAAT CAACACACAG 4500  
AGTTCCACCA TTATGCTGTG TCAAACCATA ATCCTGTATA TACAAACAAA CAAATCCAAT 4560  
CCTCTCACAG AGTCACGGTG TCGCAAAACC ACGCTAACCA TCATGGTAGC ATAGAGTAGT 4620

SUBSTITUTE SHEET (RULE 26)

- 368 -

TATTTAAAAA TTAACATAAT GATGAATTGT TAGTATGAGA TCAAAAACAA CATTGGGGCA 4680  
AATGCAACCA TGTCCAAACA CAAGAATCAA CGCACTGCCA GGA CTCTAGA AAAGACCTGG 4740  
GATACTCTTA ATCATCTAAT TGTAATATCC TCTTGT TTAT ACAGATTAAA TTTAAAATCT 4800  
ATAGCACAAA TAGCACTATC AGTTTGGCA ATGATAATCT CAACCTCTCT CATAATTGCA 4860  
GCCATAATAT TCATCATCTC TGCCAATCAC AAAGTTACAC TAACAACGGT CACAGTTCAA 4920  
ACAATAAAAA ACCACACTGA AAAAAACATC ACCACCTACC CTACTCAAGT CTCACCAGAA 4980  
AGGGTTAGTT CATCCAAGCA ACCCACAACC ACATCACCAA TCCACACAAG TTCAGCTACA 5040  
ACATCACCCA ATACAAAATC AGAAACACAC CATAACAAG CACAAACCAA AGGCAGAACC 5100  
ACCACTTCAA CACAGACCAA CAAGCCAAGC ACAAACCAC GTCCAAAAAA TCCACCAAAA 5160  
AAAGATGATT ACCATTTTGA AGTGTTCAAC TTCGTTCCCT GCAGTATATG TGGCAACAAT 5220  
CAACTTTGCA AATCCATCTG CAAAACAATA CCAAGCAACA AACCRAAGAA GAAACCAACC 5280  
ATCAAACCCA CAAACAAACC AACCACCAA ACCACAAACA AAAGAGACCC AAAAACACCA 5340  
GCCAAAACGA CGAAAAAAGA AACTACCACC AACCRAACA AAAAACTAAC CCTCAAGACC 5400  
ACAGAAAGAG ACACCAGCAC CTCACAATCC ACTGCACTCG ACACAACCAC ATTAACACAC 5460  
ACAGTCCAAC AGCAATCCCT CCTCTCAACC ACCCCCGAAA ACACACCCAA CTCCACACAA 5520  
ACACCACAG CATCCGAGCC CTCCACACCA AACTCCACCC AAAAAACCCA GCCACATGCT 5580  
TAGTTATTCA AAAACTACAT CTTAGCAGAG AACC GTGATC TATCAAGCAA GAACGAAATT 5640  
AAACCTGGGG CAAATAACCA TGGAGTTGAT GATCCACAAG TCAAGTGCAA TCTTCCTAAC 5700  
TCTTGCTATT AATGCATTGT ACCTCACCTC AAGTCAGAAC ATAAGTGGG AGTTTACCA 5760  
ATCGACATGT AGTGAGTTA GCAGAGGTTA TTTTAGTGCT TTAAGAACAG GTTGGTATAC 5820  
TAGTGTCATA ACAATAGAAT TAAGTAATAT AAAAGAAACC AAATGCAATG GAACTGACAC 5880  
TAAAGTAAAA CTTATGAAAC AAGAATTAGA TAAGTATAAG AATGCAGTAA CAGAATTACA 5940  
GCTACTTATG CAAAACACAC CAGCTGTCAA CAACCGGGCC AGAAGAGAAG CACCACAGTA 6000  
TATGAACTAC ACAATCAATA CCACTAAAAA CCTAAATGTA TCAATAAGCA AGAAGAGGAA 6060  
ACGAAGATTT CTAGGCTTCT TGTTAGGTGT GGGATCTGCA ATAGCAAGTG GTATAGCTGT 6120  
ATCAAAAGTT CTACACCTTG AAGGAGAAGT GAACAAGATC AAAAATGCTT TGTTGTCTAC 6180

SUBSTITUTE SHEET (RULE 26)



- 369 -

AAACAAAGCT GTAGTCAGTT TATCAAATGG GGTCAAGTGT TTAACCAGCA AAGTGTTAGA	6240
TCTCAAGAAT TACATAAATA ACCAATTATT ACCCATAGTA AATCAACAGA GCTGTCGCAT	6300
CTCCAACATT GAAACAGTTA TAGAATTCCA GCAGAAGAAC AGCAGATTGT TGGAAATCAC	6360
CAGAGAATTT AGTGTCAATG CAGGTGTAAC AACACCTTTA AGCACTTACA TGTTGACAAA	6420
CAGTGAGTTA CTATCATTAA TCAATGATAT GCCTATAACA AATGATCAGA AAAAATTAAT	6480
GTCAAGCAAT GTTCAGATAG TAAGGCAACA AAGTTATTCC ATCATGTCTA TAATAAAGGA	6540
AGAAGTCCTT GCATATGTTG TACAGCTGCC TATCTATGGT GTAATAGATA CACCTTGCTG	6600
GAAATTGCAC ACATCGCCTC TATGCACTAC CAACATCAAA GAAGGATCAA ATATTTGTTT	6660
AACAAGGACT GATAGAGGAT GGTATTGTGA TAATGCAGGA TCAGTATCCT TCTTTCCACA	6720
GGCTGACACT TGTAAGTAC AGTCCAATCG AGTATTTTGT GACACTATGA ACAGTTTGAC	6780
ATTACCAAGT GAAGTCAGCC TTTGTAACAC TGACATATTC AATTCCAAGT ATGACTGCAA	6840
AATTATGACA TCAAAAACAG ACATAAGCAG CTCAGTAATT ACTTCTCTTG GAGCTATAGT	6900
GTCAATGCTAT GGTAAGTAACTA AATGCACTGC ATCCAACAAA AATCGTGGGA TTATAAGAC	6960
ATTTTCTAAT GGTTGTGACT ATGTGTCAAA CAAAGGAGTA GATACTGTGT CAGTGGGCAA	7020
CACTTTTATAC TATGTAAACA AGCTGGAAGG CAAGAACCTT TATGTAAAAG GGGAACCTAT	7080
AATAAATTAC TATGACCCTC TAGTGTTCCT TTCTGATGAG TTTGATGCAT CAATATCTCA	7140
AGTCAATGAA AAAATCAATC AAAGTTTAGC TTTTATTCGT AGATCTGATG AATTACTACA	7200
TAATGTAAAT ACTGGCAAAT CTACTACAAA TATTATGATA ACTACAATTA TTATAGTAAT	7260
CATTGTAGTA TTGTTATCAT TAATAGCTAT TGGTTTACTG TTGTATTGTA AAGCCAAAAA	7320
CACACCAGTT AACTAAGCA AAGACCAACT AAGTGGAATC AATAATATTG CATTAGCAA	7380
ATAGACAAAA AACCACCTGA TCATGTTTCA ACAACAATCT GCTGACCACC AATCCCAAAT	7440
CAACTTACAA CAAATATTTT AACATCACAG TACAGGCTGA ATCATTTCCT CACATCATGC	7500
TACCCACATA ACTAAGCTAG ATCCTTAACT TATAGTTACA TAAAAACCTC AAGTATCACA	7560
ATCAACCACT AAATCAACAC ATCATTACAA AAATTAACAG CTGGGGCAAA TATGTCGCGA	7620
AGAAATCCTT GTAAATTTGA GATTAGAGGT CATTGCTTGA ATGGTAGAAG ATGTCCTAC	7680
AGTCATAATT ACTTTGAATG GCCTCCTCAT GCATTACTAG TGAGGCAAAA CTTTCATGTTA	7740

SUBSTITUTE SHEET (RULE 26)

- 370 -

AACAAGATAC TCAAGTCAAT GGACAAAAGC ATAGACACTT TGTCTGAAAT AAGTGGAGCT 7800  
GCTGAACTGG ATAGAACAGA AGAATATGCT CTTGGTATAG TTGGAGTGCT AGAGAGTTAC 7860  
ATAGGATCTA TAAACAACAT AACAAAACAA TCAGCATGTG TTGCTATGAG TAAACTTCTT 7920  
ATTGAGATCA ATAGTGATGA CATTAAAAAG CTTAGAGATA ATGAAGAACC CAATTCACCT 7980  
AAGATAAGAG TGTACAATAC TGTATATCA TACATTGAGA GCAATAGAAA AAACAACAAG 8040  
CAAACCATCC ATCTGCTCAA GAGACTACCA GCAGACGTGC TGAAGAAGAC AATAAAGAAC 8100  
ACATTAGATA TCCACAAAAG CATAACCATA AGCAATCCAA AAGAGTCAAC TGTGAATGAT 8160  
CAAAATGACC AAACCAAAAA TAATGATATT ACCGGATAAA TATCCTTGTA GTATATCATC 8220  
CATATTGATC TCAAGTGAAA GCATGGTTGC TACATTCAAT CATAAAAAACA TATTACAATT 8280  
TAACCATAAC TATTGGATA ACCACCAGCG TTTATTAAAT CATATATTG ATGAAATTCA 8340  
TTGGACACCT AAAAAGTTAT TAGATGCCAC TCAACAATTT CTCCAACATC TTAACATCCC 8400  
TGAAGATATA TATACAGTAT ATATATTAGT GTCATAATGC TTGACCATAA CGACTCTATG 8460  
TCATCCAACC ATAAAACTAT TTTGATAAGG TTATGGGACA AAATGGATCC CATTATTAAT 8520  
GGAAACTCTG CTAATGTGTA TCTAACTGAT AGTTATTTAA AAGGTGTTAT CTCTTTTTCA 8580  
GAGTGTAATG CTTTAGGGAG TTATCTTTTT AACGGCCCTT ATCTTAAAAA TGATTACACC 8640  
AACTTAATTA GTAGACAAAG CCCACTACTA GAGCATATGA ATCTTAAAAA ACTAACTATA 8700  
ACACAGTCAT TAATATCTAG ATATCATAAA GGTGAACTGA AATTAGAAGA ACCAACTTAT 8760  
TTCCAGTCAT TACTTATGAC ATATAAAAGT ATGTCCTCGT CTGAACAAAT TGCTACAACT 8820  
AACTTACTTA AAAAAATAAT ACGAAGAGCC ATAGAAATAA GTGATGTAAA GGTGTACGCC 8880  
ATCTTGAATA AACTAGGATT AAAGGAAAAG GACAGAGTTA AGCCCAACAA TAATTCAGGT 8940  
GATGAAAAC CAGTACTTAC AACCATAATT AAAGATGATA TACTTTCGGC TGTGGAAC 9000  
AATCAATCAT ATACAAATTC AGACAAAAGT CACTCAGTAA ATCAAAATAT CACTATCAAA 9060  
ACAACACTCT TGAaaaaaatt GATGTGTTCA ATGCAACATC CTCCATCATG GTTAATACAC 9120  
TGTTCAATT TATATACAAA ATTAATAAC ATATTAACAC AATATCGATC AAATGAGGTA 9180  
AAAAGTCATG GGTATATATT AATAGATAAT CAACTTTTAA GTGGTTTTCA GTTTATTTTA 9240  
AATCAATATG GTTGATCGT TTATCATAAA GGAACAAAAA AATCACAAC TACTACTTAC 9300

SUBSTITUTE SHEET (RULE 26)

- 371 -

AATCAATTTT TGACATGGAA AGACATCAGC CTTAGCAGAT TAAATGTTTG CTTAATTACT	9360
TGGATAAGTA ATTGTTTAAA TACATTAAAC AAAAGCTTAG GGCTGAGATG TGGATTCAAT	9420
AATGTTGTGT TATCACAATT ATTTCTTTAT GGAGATTGTA TACTGAAATT ATTTCATAAT	9480
GAAGGCTTCT ACATAATAAA AGAAGTAGAG GGATTTATTA TGTCTTTAAT TCTAAACATA	9540
ACAGAAGAAG ATCAATTTAA GAAACGATTT TATAATAGCA TGCTAAATAA CATCACAGAT	9600
GCAGCTATTA AGGCTCAAAA GGACCTACTA TCAAGAGTAT GTCACACTTT ATTAGACAAG	9660
ACAGTGTCTG ATAATATCAT AAATGGTAAA TGGATAATCC TATTAAGTAA ATTTCTTAAA	9720
TTGATTAAGC TTGCAGGTGA TAATAATCTC AATAACTTGA GTGAGCTATA TTTTCTCTTC	9780
AGAATCTTTG GACATCCAAT GGTGATGAA AGACAAGCAA TGGATTCTGT AAGAATTAAC	9840
TGTAATGAAA CTAAGTTCTA CTTATTAAGT AGTCTAAGTA CATTAGAGG TGCTTTCATT	9900
TATAGAATCA TAAAAGGGTT TGTAATACC TACAACAGAT GGCCACCTT AAGGAATGCT	9960
ATTGTCCTAC CTCTAAGATG GTTAACTAC TATAAECTTA ATACTTATCC ATCTCTACTT	10020
GAAATCACAG AAAATGATT GATTATTTTA TCAGGATTGC GGTTCTATCG TGAGTTTCAT	10080
CTGCCTAAAA AAGTGGATCT TGAAATGATA ATAAATGACA AAGCCATTTC ACCTCCAAAA	10140
GATCTAATAT GGACTAGTTT TCCTAGAAAT TACATGCCAT CACATATACA AAATTATATA	10200
GAACATGAAA AGTTGAAGTT CTCTGAAAGC GACAGATCGA GAAGAGTACT AGAGTATTAC	10260
TTGAGAGATA ATAAATTCAA TGAATGCGAT CTATACAATT GTGTAGTCAA TCAAAGCTAT	10320
CTCAACAACCT CTAATCACGT GGTATCACTA ACTGGTAAAG AAAGAGAGCT CAGTGTAGGT	10380
AGAATGTTTG CTATGCAACC AGGTATGTTT AGGCAAATCC AAATCTTAGC AGAGAAAATG	10440
ATAGCTGAAA ATATTTTACA ATTCTTCCCT GAGAGTTTGA CAAGATATGG TGATCTAGAG	10500
CTTCAAAAAGA TATTAGAATT AAAAGCAGGA ATAAGCAACA AGTCAAATCG TTATAATGAT	10560
AACTACAACA ATTATATCAG TAAATGTTCT ATCATTACAG ATCTTAGCAA ATTCAATCAG	10620
GCATTTAGAT ATGAAACATC ATGTATCTGC AGTGATGTAT TAGATGAACT GCATGGAGTA	10680
CAATCTCTGT TCTCTTGGTT GCATTTAACA ATACCTCTTG TCACAATAAT ATGTACATAT	10740
AGACATGCAC CTCCTTTTCAT AAAGGATCAT GTTGTTAATC TTAATGAGGT TGATGAACAA	10800
AGTGGATTAT ACAGATATCA TATGGGTGGT ATTGAGGGCT GGTGTCAAAA ACTGTGGACC	10860

SUBSTITUTE SHEET (RULE 26)

- 372 -

ATTGAAGCTA TATCATTATT AGATCTAATA TCTCTCAAAG GGAAATTCTC TATCACAGCT 10920  
CTGATAAATG GTGATAATCA GTCAATTGAT ATAAGCAAAC CAGTTAGACT TATAGAGGGT 10980  
CAGACCCATG CACAAGCAGA TTATTTGTGA GCATTAAATA GCCTTAAATT GTTATATAAA 11040  
GAGTATGCAG GTATAGGCCA TAAGCTTAAG GGAACAGAGA CCTATATATC CCGAGATATG 11100  
CAGTTCATGA GCAAAACAAT CCAGCACAAT GGAGTGTACT ATCCAGCCAG TATCAAAAAA 11160  
GTCCTGAGAG TAGGTCCATG GATAAACACG ATACTTGATG ATTTTAAAGT TAGTTTAGAA 11220  
TCTATAGGCA GCTTAACACA GGAGTTAGAA TACAGAGGAG AAAGCTTATT ATGCAGTTTA 11280  
ATATTTAGGA ACATTTGGTT ATACAATCAA ATTGCTTTC AACTCCGAAA TCATGCATTA 11340  
TGTAACAATA AGCTATATTT AGATATATTG AAAGTATTAA AACACTTAAA AACTTTTTTT 11400  
AATCTTGATA GCATTGATAT GGCTTTATCA TTGTATATGA ATTTGCCTAT GCTGTTTGGT 11460  
GGTGGTGATC CTAATTTGTT ATATCGAAGC TTTTATAGGA GAACTCCAGA CTTCCTTACA 11520  
GAAGCTATAG TACATTCAGT GTTTGTGTTG AGCTATTATA CTGGTCACGA TTTACAAGAT 11580  
AAGCTCCAGG ATCTTCAGA TGATAGACTG AACAAATTCT TGACATGTGT CATCACATTT 11640  
GATAAAATC CCAATGCCGA GTTTGTAACA TTGATGAGGG ATCCACAGGC TTTAGGGTCT 11700  
GAAAGGCAAG CTAATAATTAC TAGTGAGATT AATAGATTAG CAGTAACAGA AGTCTTAAGT 11760  
ATAGCCCCAA ACAAATATT TTCTAAAAGT GCACAACATT ATACTACCAC TGAGATTGAT 11820  
CTAAATGACA TTATGCAAAA TATAGAACCA ACTTACCCTC ATGGATTAAG AGTTGTTTAT 11880  
GAAAGTTTAC CTTTTTATAA AGCAGAAAAA ATAGTTAATC TTATATCAGG AACAAAATCC 11940  
ATAACTAATA TACTTGAAAA AACATCAGCA ATAGATACAA CTGATATTAA TAGGGCTACT 12000  
GATATGATGA GGAAAAATAT AACTTTACTT ATAAGGATAC TTCCACTAGA TTGTAACAAA 12060  
GACAAAAGAG AGTTATTAAG TTTAGAAAAT CTTAGTATAA CTGAATTAAG CAAGTATGTA 12120  
AGAGAAAGAT CTTGGTCATT ATCCAATATA GTAGGAGTAA CATCGCCAAG TATTATGTTT 12180  
ACAATGAACA TTAAATATAC AACTAGCACT ATAGCCAGTG GTATAATAAT AGAAAAATAT 12240  
AATGTTAATA GTTTAACTCG TGGTGAAAGA GGACCCACCA AGCCATGGGT AGGCTCATCC 12300  
ACGCAGGAGA AAAAAACAAT GCCAGTGATC AACAGACAAG TTTTAACCAA AAAGCAAAGA 12360  
GACCAAATAG ATTTATTAGC AAAATTAGAC TGGGTATATG CATCCATAGA CAACAAAGAT 12420

SUBSTITUTE SHEET (RULE 26)

- 373 -

GAATTCATGG AAGAACTGAG TACTGGAACA CTGGGACTGT CATATGAAAA AGCCAAAAAG	12480
TTGTTTCCAC AATATCTAAG TGTCAATTAT TTACACCGTT TAACAGTCAG TAGTAGACCA	12540
TGTGAATTCC CTGCATCAAT ACCAGCTTAT AGAACAACAA ATTATCATT TGATACTAGT	12600
CCTATCAATC ATGTATTAAC AGAAAAGTAT GGAGATGAAG ATATCGACAT TGTGTTTCAA	12660
AATTGCATAA GTTTTGGTCT TAGCCTGATG TCGGTTGTGG AACAAATTCAC AAACATATGT	12720
CCTAATAGAA TTATTCTCAT ACCGAAGCTG AATGAGATAC ATTTGATGAA ACCTCCTATA	12780
TTTACAGGAG ATGTTGATAT CATCAAGTTG AAGCAAGTGA TACAAAAGCA GCACATGTTT	12840
CTACCAGATA AAATAAGTTT AACCCAATAT GTAGAATTAT TCTTAAGTAA CAAAGCACTT	12900
AAATCTGGAT CTCACATCAA CTCTAATTTA ATATTAGTAC ATAAAATGTC TGATTATTTT	12960
CATAATGCTT ATATTTTAAG TACTAATTTA GCTGGACATT GGATTCTGAT TATTCAACTT	13020
ATGAAAGATT CAAAAGGTAT TTTTGAAAAA GATTGGGGAG AGGGGTACAT AACTGATCAT	13080
ATGTTCATTA ATTTGAATGT TTTCTTTAAT GCTTATAAGA CTTATTTGCT ATGTTTTCAT	13140
AAAGGTTATG GTAAAGCAAA ATTAGAATGT GATATGAACA CTTGAGATCT TCTTTGTGTT	13200
TTGGAGTTAA TAGACAGTAG CTACTGGAAA TCTATGTCTA AAGTTTTCCT AGAACAAAAA	13260
GTCATAAAAT ACATAGTCAA TCAAGACACA AGTTTGCGTA GAATAAAAGG CTGTCACAGT	13320
TTTAAGTTGT GGTTTTTTAAA ACGCCTTAAT AATGCTAAAT TTACCGTATG CCCTTGGGTT	13380
GTTAACATAG ATTATCACCC AACACACATG AAAGCTATAT TATCTTACAT AGATTTAGTT	13440
AGAATGGGGT TAATAAATGT AGATAAATTA ACCATTAAAA ATAAAAACAA ATTCAATGAT	13500
GAATTTTACA CATCAAATCT CTTTTACATT AGTTATAACT TTTGAGACAA CACTCATTTG	13560
CTAACAAAAC AAATAAGAAT TGCTAATTCA GAATTAGAAG ATAATTATAA CAACTATAT	13620
CACCCAACCC CAGAACTTT AGAAAATATG TCATTAATTC CTGTTAAAAG TAATAATAGT	13680
AACAAACCTA AATTTTGTAT AAGTGGAAT ACCGAATCTA TGATGATGTC AACATTCTCT	13740
AGTAAAATGC ATATTAAATC TTCCACTGTT ACCACAAGAT TCAATTATAG CAAACAAGAC	13800
TTGTACAATT TATTTCCAAT TGTGTGATA GACAAGATTA TAGATCATTC AGGTAATACA	13860
GCAAAATCTA ACCAACTTTA CACCACCACT TCACATCAGA CATCTTTAGT AAGGAATAGT	13920
GCATCACTTT ATTGCATGCT TCCTTGGCAT CATGTCAATA GATTAACTT TGTATTTAGT	13980

SUBSTITUTE SHEET (RULE 26)

- 374 -

TCCACAGGAT GCAAGATCAG TATAGAGTAT ATTTTAAAAG ATCTTAAGAT TAAGGACCCC 14040  
AGTTGTATAG CATTCATAGG TGAAGGAGCT GGTAACCTAT TATTACGTAC GGTAGTAGAA 14100  
CTTCATCCAG ACATAAGATA CATTACAGA AGTTTAAAAG ATTGCAATGA TCATAGTTTA 14160  
CCTATTGAAT TTCTAAGGTT ATACAACGGG CATATAACA TAGATTATGG TGAGAATTTA 14220  
ACCATTCCCTG CTACAGATGC AACTAATAAC ATTCATTGGT CTTATTTACA TATAAAATTT 14280  
GCAGAACCTA TTAGCATCTT TGTCTGCGAT GCTGAATTAC CTGTTACAGC CAATTGGAGT 14340  
AAAATTATAA TTGAATGGAG TAAGCATGTA AGAAAGTGCA AGTACTGTTT TTCTGTAAAT 14400  
AGATGCATTT TAATTGCAAA ATATCATGCT CAAGATGACA TTGATTTCAA ATTAGATAAC 14460  
ATTACTATAT TAAAACTTA CGTGTGCCTA GGTAGCAAGT TAAAAGGATC TGAAGTTTAC 14520  
TTAATCCTTA CAATAGGCCC TGCAAATATA CTTCTGTTT TTGATGTTGT ACAAATGCT 14580  
AAATTGATAC TTTCAAGAAC TAAAAATTTT ATTATGCCTA AAAAACTGA CAAGGAATCT 14640  
ATCGATGCAA ATATTAAAAG CTTAATACCT TTCCTTTGTT ACCCTATAAC AAAAAAGGA 14700  
ATTAAGACTT CATTGTCAA ATTGAAGAGT GTAGTTAATG GAGATATATT ATCATATTC 14760  
ATAGCTGGAC GTAATGAAGT ATTCAGCAAC AAGCTTATAA ACCACAAGCA TATGAATATC 14820  
CTAAAATGGC TAGATCATGT TTTAAATTTT AGATCAGCTG AACTTAATTA CAATCATTTA 14880  
TACATGATAG AGTCCACATA TCCTTACTTA AGTGAATTGT TAAATAGTTT AACACCAAT 14940  
GAGCTCAAGA AGCTGATTAA AATAACAGGT AGTGTGCTAT ACAACCTTCC CAACGAACAG 15000  
TAGTTTAAAA TATCATTAAC AAGTTTGGTC AAATTTAGAT GCTAACACAT CATTATATTA 15060  
TAGTTATTAA AGAATATACA AACTTTTCAA TAATTTAGCA TATTGATTCC AAAATTATCA 15120  
TTTGTAGTCTT AAGGGGTTAA ATAAAAGTCT AAAACTAACA ATTATACATG TGCATTCACA 15180  
ACACAACGAG ACATTAGTTT TTGACACTTT TTTTCTCGT 15219

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

SUBSTITUTE SHEET (RULE 26)

- 375 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Met Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp  
 1 5 10 15  
 Ser Tyr Leu Lys Gly Val Ile Ser Phe Ser Glu Cys Asn Ala Leu Gly  
 20 25 30  
 Ser Tyr Leu Phe Asn Gly Pro Tyr Leu Lys Asn Asp Tyr Thr Asn Leu  
 35 40 45  
 Ile Ser Arg Gln Ser Pro Leu Leu Glu His Met Asn Leu Lys Lys Leu  
 50 55 60  
 Thr Ile Thr Gln Ser Leu Ile Ser Arg Tyr His Lys Gly Glu Leu Lys  
 65 70 75 80  
 Leu Glu Glu Pro Thr Tyr Phe Gln Ser Leu Leu Met Thr Tyr Lys Ser  
 85 90 95  
 Met Ser Ser Ser Glu Gln Ile Ala Thr Thr Asn Leu Leu Lys Lys Ile  
 100 105 110  
 Ile Arg Arg Ala Ile Glu Ile Ser Asp Val Lys Val Tyr Ala Ile Leu  
 115 120 125  
 Asn Lys Leu Gly Leu Lys Glu Lys Asp Arg Val Lys Pro Asn Asn Asn  
 130 135 140  
 Ser Gly Asp Glu Asn Ser Val Leu Thr Thr Ile Ile Lys Asp Asp Ile  
 145 150 155 160  
 Leu Ser Ala Val Glu Asn Asn Gln Ser Tyr Thr Asn Ser Asp Lys Ser  
 165 170 175  
 His Ser Val Asn Gln Asn Ile Thr Ile Lys Thr Thr Leu Leu Lys Lys  
 180 185 190  
 Leu Met Cys Ser Met Gln His Pro Pro Ser Trp Leu Ile His Trp Phe  
 195 200 205  
 Asn Leu Tyr Thr Lys Leu Asn Asn Ile Leu Thr Gln Tyr Arg Ser Asn  
 210 215 220  
 Glu Val Lys Ser His Gly Phe Ile Leu Ile Asp Asn Gln Thr Leu Ser  
 225 230 235 240  
 Gly Phe Gln Phe Ile Leu Asn Gln Tyr Gly Cys Ile Val Tyr His Lys

SUBSTITUTE SHEET (RULE 26)

- 376 -

245	250	255
Gly Leu Lys Lys Ile Thr Thr Thr Thr Tyr Asn Gln Phe Leu Thr Trp		
260	265	270
Lys Asp Ile Ser Leu Ser Arg Leu Asn Val Cys Leu Ile Thr Trp Ile		
275	280	285
Ser Asn Cys Leu Asn Thr Leu Asn Lys Ser Leu Gly Leu Arg Cys Gly		
290	295	300
Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile		
305	310	315
Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu		
325	330	335
Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe		
340	345	350
Lys Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala		
355	360	365
Ile Lys Ala Gln Lys Asp Leu Leu Ser Arg Val Cys His Thr Leu Leu		
370	375	380
Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu		
385	390	395
Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu		
405	410	415
Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro		
420	425	430
Met Val Asp Glu Arg Gln Ala Met Asp Ser Val Arg Ile Asn Cys Asn		
435	440	445
Glu Thr Lys Phe Tyr Leu Leu Ser Ser Leu Ser Thr Leu Arg Gly Ala		
450	455	460
Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp		
465	470	475
Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr		
485	490	495
Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Asn Asp		
500	505	510
Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro		
515	520	525

SUBSTITUTE SHEET (RULE 26)



- 377 -

Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro  
 530 535 540

Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser  
 545 550 555 560

His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser  
 565 570 575

Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe  
 580 585 590

Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn  
 595 600 605

Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser  
 610 615 620

Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln  
 625 630 635 640

Ile Leu Ala Glu Lys Met Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro  
 645 650 655

Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu  
 660 665 670

Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr  
 675 680 685

Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe  
 690 695 700

Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu  
 705 710 715 720

Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr  
 725 730 735

Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe  
 740 745 750

Ile Lys Asp His Val Val Asn Leu Asn Glu Val Asp Glu Gln Ser Gly  
 755 760 765

Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu  
 770 775 780

Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly  
 785 790 795 800

SUBSTITUTE SHEET (RULE 26)

Lys	Phe	Ser	Ile	Thr	Ala	Leu	Ile	Asn	Gly	Asp	Asn	Gln	Ser	Ile	Asp	
																805
																810
																815
Ile	Ser	Lys	Pro	Val	Arg	Leu	Ile	Glu	Gly	Gln	Thr	His	Ala	Gln	Ala	
																820
																825
																830
Asp	Tyr	Leu	Leu	Ala	Leu	Asn	Ser	Leu	Lys	Leu	Leu	Tyr	Lys	Glu	Tyr	
																835
																840
																845
Ala	Gly	Ile	Gly	His	Lys	Leu	Lys	Gly	Thr	Glu	Thr	Tyr	Ile	Ser	Arg	
																850
																855
																860
Asp	Met	Gln	Phe	Met	Ser	Lys	Thr	Ile	Gln	His	Asn	Gly	Val	Tyr	Tyr	
																865
																870
																875
																880
Pro	Ala	Ser	Ile	Lys	Lys	Val	Leu	Arg	Val	Gly	Pro	Trp	Ile	Asn	Thr	
																885
																890
																895
Ile	Leu	Asp	Asp	Phe	Lys	Val	Ser	Leu	Glu	Ser	Ile	Gly	Ser	Leu	Thr	
																900
																905
																910
Gln	Glu	Leu	Glu	Tyr	Arg	Gly	Glu	Ser	Leu	Leu	Cys	Ser	Leu	Ile	Phe	
																915
																920
																925
Arg	Asn	Ile	Trp	Leu	Tyr	Asn	Gln	Ile	Ala	Leu	Gln	Leu	Arg	Asn	His	
																930
																935
																940
Ala	Leu	Cys	Asn	Asn	Lys	Leu	Tyr	Leu	Asp	Ile	Leu	Lys	Val	Leu	Lys	
																945
																950
																955
His	Leu	Lys	Thr	Phe	Phe	Asn	Leu	Asp	Ser	Ile	Asp	Met	Ala	Leu	Ser	
																965
																970
																975
Leu	Tyr	Met	Asn	Leu	Pro	Met	Leu	Phe	Gly	Gly	Gly	Asp	Pro	Asn	Leu	
</																

BNSDOCID: <WO 9813501A2>

- 379 -

1075	1080	1085
Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu 1090	1095	1100
Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His 1105	1110	1115 1120
Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys 1125	1130	1135
Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu 1140	1145	1150
Lys Thr Ser Ala Ile Asp Thr Thr Asp Ile Asn Arg Ala Thr Asp Met 1155	1160	1165
Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys 1170	1175	1180
Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr 1185	1190	1195 1200
Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile 1205	1210	1215
Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asn Ile Lys Tyr 1220	1225	1230
Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val 1235	1240	1245
Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly 1250	1255	1260
Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val 1265	1270	1275 1280
Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp 1285	1290	1295
Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu 1300	1305	1310
Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe 1315	1320	1325
Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser 1330	1335	1340
Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn 1345	1350	1355 1360

SUBSTITUTE SHEET (RULE 26)

- 380 -

Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr  
 1365 1370 1375  
 Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly  
 1380 1385 1390  
 Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn  
 1395 1400 1405  
 Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro  
 1410 1415 1420  
 Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile  
 1425 1430 1435 1440  
 Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr  
 1445 1450 1455  
 Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile  
 1460 1465 1470  
 Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn  
 1475 1480 1485  
 Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile  
 1490 1495 1500  
 Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu  
 1505 1510 1515 1520  
 Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn  
 1525 1530 1535  
 Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala  
 1540 1545 1550  
 Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu  
 1555 1560 1565  
 Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu  
 1570 1575 1580  
 Gln Lys Val Ile Lys Tyr Ile Val Asn Gln Asp Thr Ser Leu Arg Arg  
 1585 1590 1595 1600  
 Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asn  
 1605 1610 1615  
 Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His  
 1620 1625 1630

SUBSTITUTE SHEET (RULE 26)

- 381 -

Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met  
 1635 1640 1645  
 Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe  
 1650 1655 1660  
 Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe  
 1665 1670 1675 1680  
 Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser  
 1685 1690 1695  
 Glu Leu Glu Asp Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr  
 1700 1705 1710  
 Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys  
 1715 1720 1725  
 Pro Lys Phe Cys Ile Ser Gly Asn Thr Glu Ser Met Met Met Ser Thr  
 1730 1735 1740  
 Phe Ser Ser Lys Met His Ile Lys Ser Ser Thr Val Thr Thr Arg Phe  
 1745 1750 1755 1760  
 Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile  
 1765 1770 1775  
 Asp Lys Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu  
 1780 1785 1790  
 Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser  
 1795 1800 1805  
 Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val  
 1810 1815 1820  
 Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp  
 1825 1830 1835 1840  
 Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala  
 1845 1850 1855  
 Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg  
 1860 1865 1870  
 Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile  
 1875 1880 1885  
 Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu  
 1890 1895 1900  
 Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser

SUBSTITUTE SHEET (RULE 26)

- 382 -

1905	1910	1915	1920
Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp			
	1925	1930	1935
Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp			
	1940	1945	1950
Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys			
	1955	1960	1965
Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu			
	1970	1975	1980
Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu			
	1985	1990	2000
Lys Gly Ser Glu Val Tyr Leu Ile Leu Thr Ile Gly Pro Ala Asn Ile			
	2005	2010	2015
Leu Pro Val Phe Asp Val Val Gln Asn Ala Lys Leu Ile Leu Ser Arg			
	2020	2025	2030
Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp			
	2035	2040	2045
Ala Asn Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys			
	2050	2055	2060
Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Asn Gly			
	2065	2070	2075
Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn			
	2085	2090	2095
Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His			
	2100	2105	2110
Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met			
	2115	2120	2125
Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr			
	2130	2135	2140
Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr			
	2145	2150	2155
Asn Leu Pro Asn Glu Gln			
	2165		

(2) INFORMATION FOR SEQ ID NO:33:

SUBSTITUTE SHEET (RULE 26)

- 383 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15219 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ACGGGAAAAA AATGCGTACT ACAAACCTTGC ACATTCGAAA AAAATGGGGC AAATAAGAAC	60
TTGATAAGTG CTATTTAAGT CTAACCTTTT CAATCAGAAA TGGGGTGCAA TTTACTGAGC	120
ATGATAAAGG TTAGATTACA AAATTTATTT GACAATGACG AAGTAGCATT GTTAAAAATA	180
ACATGTTATA CTGATAAATT AATTCTTCTG ACCAATGCAT TAGCCAAAGC AGCAATACAT	240
ACAATTAAAT TAAACGGCAT AGTTTTTATA CATGTTATAA CAAGCAGTGA AGTGTGCCCT	300
GATAACAATA TTGTAGTGAA ATCTAACTTT ACAACAATGC CAATACTACA AAATGGAGGA	360
TACATATGGG AATTGATTGA GTTGACACAC TGCTCTCAAT TAAACGGTTT AATGGATGAT	420
AATTGTGAAA TCAAATTTTC TAAAAGACTA AGTGACTCAG TAATGACTAA TTATATGAAT	480
CAAATATCTG ACTTACTTGG GCTTGATCTC AATTCATGAA TTATGTTTAG TCTAATTCAA	540
TAGACATGTG TTTATTACCA TTTTAGTTAA TATAAAAACT CATCAAAGGG AAATGGGGCA	600
AATAAACTCA CCTAATCAAT CAAACCATGA GCACTACAAA TGACAACACT ACTATGCAAA	660
GATTGATGAT CACAGACATG AGACCCCTGT CAATGGATTC AATAATAACA TCTCTTACCA	720
AAGAAATCAT CACACACAAA TTCATATACT TGATAAACAA TGAATGTATT GTAAGAAAAC	780
TTGATGAAAG ACAAGCTACA TTTACATTCT TAGTCAATTA TGAGATGAAG CTACTGCACA	840
AAGTAGGGAG TACCAAATAC AAAAAATACA CTGAATATAA TACAAAATAT GGCACCTTCC	900
CCATGCCTAT ATTTATCAAT CACGGCGGGT TTCTAGAATG TATTGGCATT AAGCCTACAA	960
AACACACTCC TATAATATAC AAATATGACC TCAACCCGTG AATTCCAACA AAAAAACCAA	1020
CCCAACCAAA CCAAACCTATT CCTCAAACAA CAGTGCTCAA TAGTTAAGAA GGAGCTAATC	1080
CATTTTAGTA ATTA AAAATA AAAGTAAAGC CAATAACATA AATTGGGGCA AATACAAAGA	1140
TGGCTCTTAG CAAAGTCAAG TTGAATGATA CATTAAATAA GGATCAGCTG CTGTCATCCA	1200

SUBSTITUTE SHEET (RULE 26)

- 384 -

GCAAATACAC TATTCAACGT AGTACAGGAG ATAATATTGA CACTCCCAAT TATGATGTGC 1260  
AAAAACACCT AAACAAACTA TGTGGTATGC TATTAATCAC TGAAGATGCA AATCATAAAT 1320  
TCACAGGATT AATAGGTATG TTATATGCTA TGTCCAGGTT AGGAAGGGAA GACACTATAA 1380  
AGATACTTAA AGATGCTGGA TATCATGTTA AAGCTAATGG AGTAGATATA ACAACATATC 1440  
GTCAAGATAT AAATGGAAAG GAAATGAAAT TCGAAGTATT AACATTATCA AGCTTGACAT 1500  
CAGAAATACA AGTCAATATT GAGATAGAAT CTAGAAAGTC CTACAAAAAA ATGCTAAAAG 1560  
AGATGGGAGA AGTGGCTCCA GAATATAGGC ATGATTCTCC AGACTGTGGG ATGATAATAC 1620  
TGTGTATAGC TGCACTTGTG ATAACCAAAT TAGCAGCAGG AGACAGATCA GGTCTTACAG 1680  
CAGTAATTAG GAGGGCAAAC AATGTCTTAA AAAACGAAAT AAAACGATAC AAGGGCCTCA 1740  
TACCAAAGGA TATAGCTAAC AGTTTTTATG AAGTGTTTGA AAAACACCCT CATCTTATAG 1800  
ATGTTTTCGT GCACTTTGGC ATTGCACAAT CATCCACAAG AGGGGGTAGT AGAGTTGAAG 1860  
GAATCTTTGC AGGATTGTTT ATGAATGCCT ATGGTTCAGG GCAAGTAATG CTAAGATGGG 1920  
GAGTTTTAGC CAAATCTGTA AAAAATATCA TGCTAGGACA TGCTAGTGTC CAGGCAGAAA 1980  
TGGAGCAAGT TGTGGAAGTC TATGAGTATG CACAGAAGTT GGGAGGAGAA GCTGGATTCT 2040  
ACCATATATT GAACAATCCA AAAGCATCAT TGCTGTCATT AACTCAATTT CCCAACTTCT 2100  
CAAGTGTGGT CCTAGGCAAT GCAGCAGGTC TAGGCATAAT GGGAGAGTAT AGAGGTACAC 2160  
CAAGAAACCA GGATCTTTAT GATGCAGCTA AAGCATATGC AGAGCAACTC AAAGAAAATG 2220  
GAGTAATAAA CTACAGTGTA TTAGACTTAA CAGCAGAAGA ATTGGAAGCC ATAAAGCATC 2280  
AACTCAACCC CAAAGAAGAT GATGTAGAGC TTAAAGTTAA CAAAAAATAC GGGGCAAATA 2340  
AGTCAACATG GAGAAGTTG CACCTGAATT TCATGGAGAA GATGCAAATA ACAAAGCTAC 2400  
CAAATTCCTA GAATCAATAA AGGGCAAGTT CGCATCATCC AAAGATCCTA AGAAGAAAGA 2460  
TAGCATAATA TCTGTTAACT CAATAGATAT AGAAGTAACT AAAGAGAGCC CGATAACATC 2520  
TGGCACCAAC ATCATCAATC CAACAAGTGA AGCCGACAGT ACCCCAGAAA CAAAAGCCAA 2580  
CTACCCAAGA AAACCCCTAG TAAGCTTCAA AGAAGATCTC ACCCCAAGTG ACAACCCTTT 2640  
TTCTAAGTTG TACAAGGAAA CAATAGAAAC ATTTGATAAC AATGAAGAAG AATCTAGCTA 2700  
CTCATATGAA GAGATAAATG ATCAAACAAA TGACAACATT ACAGCAAGAC TAGATAGAAT 2760

SUBSTITUTE SHEET (RULE 26)



- 385 -

TGATGAAAAA TTAAGTGAAA TATTAGGAAT GCTCCATACA TTAGTAGTTG CAAGTGCAGG	2820
ACCCACTTCA GCTCGCGATG GAATAAGAGA TGCTATGGTT GGTCTAAGAG AAGAGATGAT	2880
AGAAAAAATA AGAGCGGAAG CATTAAATGAC CAATGATAGG TTAGAGGCTA TGGCAAGACT	2940
TAGGAATGAG GAAAGCGAAA AAATGGCAAA AGACACCTCA GATGAAGTGT CTCTTAATCC	3000
AACTTCCAAA AAATTGAGTG ACTTGTTGGA AGACAACGAT AGTGACAATG ATCTATCACT	3060
TGATGATTTT TGATCAGCGA TCAACTCACT CAGCAATCAA CAACATCAAT AAAACAGACA	3120
TCAATCCATT GAATCAACTG CCAGACCGAA CAAACAAACG TCCATCAGTA GAACCACCAA	3180
CCAATCAATC AACCAATTGA TCAATCAGCA ACCCGACAAA ATTAACAATA TAGTAACAAA	3240
AAAAGAACAA GATGGGGCAA ATATGGAAAC ATACGTGAAC AAGCTTCACG AAGGCTCCAC	3300
ATACACAGCA GCTGTTTCAGT ACAATGTTCT AGAAAAAGAT GATGATCCTG CATCACTAAC	3360
AATATGGGTG CCTATGTTCC AGTCATCTGT GCCAGCAGAC TTGCTCATAA AAGAACTTGC	3420
AAGCATCAAT ATACTAGTGA AGCAGATCTC TACGCCCCAA GGACCTTCAC TACGAGTCAC	3480
GATTAECTCA AGAAGTGCTG TGCTGGCTCA AATGCCTAGT AATTTTCATCA TAAGCGCAAA	3540
TGTATCATTG GATGAAAGAA GCAAATTAGC ATATGATGTA ACTACACCTT GTGAAATCAA	3600
AGCATGCAGT CTAACATGCT TAAAAGTAAA AAGTATGTTA ACTACAGTCA AAGATCTTAC	3660
CATGAAGACA TTCAACCCCA CTCATGAGAT CATTGCTCTA TGTGAATTTG AAAATATTAT	3720
GACATCAAAA AGAGTAATAA TACCAACCTA TCTAAGATCA ATTAGTGTC AAGAACAGGA	3780
TCTGAACTCA CTAGAAAATA TAGCAACCAC CGAATTCAAA AATGCTATCA CCAATGCAAA	3840
AATTATTCCT TATGCAGGAT TAGTGTTAGT TATCACAGTT ACTGACAATA AAGGAGCATT	3900
CAAATATATC AAACCACAGA GTCAATTTAT AGTAGATCTT GGTGCCTACC TAGAAAAAGA	3960
GAGCATATAT TATGTGACTA CTAATTGGAA GCATACAGCT ACACGTTTTT CAATCAAACC	4020
ACTAGAGGAT TAACTTAAT TATCAACACT GAATGACAGG TCCACATATA TCCTCAAAC	4080
ACACACTATA TCCAAACATC ATAAACATCT AACTACACA CTTTCATCACA CAAACCAATC	4140
CCACTCAAAA TCCAAAATCA CTACCAGCCA CTATCTGCTA GACCTAGAGT GCGAATAGGT	4200
AAATAAAACC AAAATATGGG GTAAATAGAC ATTAGTTAGA GTTCAATCAA TCTTAACAAC	4260
CATTATATACC GCCAATTCAA CACATATACT ATAAATCTTA AAATGGGAAA TACATCCATC	4320

SUBSTITUTE SHEET (RULE 26)

- 386 -

ACAATAGAAT TCACAAGCAA ATTTTGGCCC TATTTTACAC TAATACATAT GATCTTAACT	4380
CTAATCTTTT TACTAATTAT AATCACTATT ATGATTGCAA TACTAAATAA GCTAAGTGAA	4440
CATAAAGCAT TCTGTAACAA AACTCTTGAA CTAGGACAGA TGTATCAAAT CAACACATAG	4500
AGTTCTACCA TTATGCTGTG TCAAATTATA ATCCTGTATA TATAAACAAA CAAATCCAAT	4560
CTTCTCACAG AGTCATGGTG TCGCAAAACC ACGCTAACTA TCATGGTAGC ATAGAGTAGT	4620
TATTTAAAAA TTAACATAAT GATGAATTGT TAGTATGAGA TCAAAAACAA CATTGGGGCA	4680
AATGCAACCA TGTCCAAACA CAAGAATCAA CGCACTGCCA GGACTCTAGA AAAGACCTGG	4740
GATACTCTTA ATCATCTAAT TGTAATATCC TCTTGTTTAT ACAGATTAAA TTTAAATCT	4800
ATAGCACAAA TAGCACTATC AGTTTTGGCA ATGATAATCT CAACCTCTCT CATAATTGCA	4860
GCCATAATAT TCATCATCTC TGCCAATCAC AAAGTTACAC TAACAACGGT CACAGTTCAA	4920
ACAATAAAAA ACCCACTGA AAAAAACATC ACCACCTACC CTAICTAAGT CTCACCAGAA	4980
AGGGTTAGTT CATCCAAGCA ACCCACAACC ACATCACCAA TCCACACAAG TTCAGCTACA	5040
ACATCACCCA ATACAAAATC AGAAACACAC CATAACACAG CACAAACCAA AGGCAGAACC	5100
ACCACTTCAA CACAGACCAA CAAGCCAAGC ACAAACCCAC GTCCAAAAAA TCCACCAAAA	5160
AAAGATGATT ACCATTTTGA AGTGTTCAC TTCGTTCCCT GCAGTATATG TGGCAACAAT	5220
CAACTTTGCA AATCCATCTG CAAAACAATA CCAAGCAACA AACCAGAGAA GAAACCAACC	5280
ATCAAACCCA CAAACAAACC AACCACCAAA ACCACAAACA AAAGAGACCC AAAACACCA	5340
GCCAAAACGA CGAAAAAGA AACTACCACC AACCACAA AAAAACTAAC CCTCAAGACC	5400
ACAGAAAGAG ACACCAGCAC CTCACAATCC ACTGCACTCG ACACAACCAC ATTAACACAC	5460
ACAGTCCAAC AGCAATCCCT CCTCTCAACC ACCCCCGAAA ACACACCCAA CTCCACACAA	5520
ACACCCACAG CATCCGAGCC CTCCACACCA AACTCCACCC AAAAAACCCA GCCACATGCT	5580
TAGTTATTCA AAAACTACAT CTTAGCAGAG AACCCTGATC TATCAAGCAA GAACGAAATT	5640
AAACCTGGGG CAAATAACCA TGGAGTTGAT GATCCACAAG TCAAGTGCAA TCTTCCTAAC	5700
TCTTGCTATT AATGCATTGT ACCTCACCTC AAGTCAGAAC ATAAGTGAGG AGTTTTACCA	5760
ATCGACATGT AGTGCAGTTA GCAGAGGTTA TTTTAGTGCT TTAAGAACAG GTTGGTATAC	5820
TAGTGTGATA ACAATAGAAT TAAGTAATAT AAAAGAAACC AAATGCAATG GAACTGACAC	5880

SUBSTITUTE SHEET (RULE 26)

- 387 -

TAAAGTAAAA CTTATGAAAC AAGAATTAGA TAAGTATAAG AATGCAGTAA CAGAATTACA 5940  
GCTACTTATG CAAAACACAC CAGCTGTCAA CAACCGGGCC AGAAGAGAAG CACCACAGTA 6000  
TATGAACTAC ACAATCAATA CCACTAAAAA CCTAAATGTA TCAATAAGCA AGAAGAGGAA 6060  
ACGAAGATTT CTAGGCTTCT TGTTAGGTGT GGGATCTGCA ATAGCAAGTG GTATAGCTGT 6120  
ATCAAAAGTT CTACACCTTG AAGGAGAAGT GAACAAGATC AAAAATGCTT TGTGTCTAC 6180  
AAACAAAGCT GTAGTCAGTT TATCAAATGG GGTCAAGTGT TTAACCAGCA AAGTGTTAGA 6240  
TCTCAAGAAT TACATAAATA ACCAATTATT ACCCATAGTA AATCAACAGA GCTGTCGCAT 6300  
CTCCAACATT GAAACAGTTA TAGAATTCCA GCAGAAGAAC AGCAGATTGT TGGAAATCAC 6360  
CAGAGAATTT AGTGTCAATG CAGGTGTAAC AACACCTTTA AGCACTTACA TGTGACAAA 6420  
CAGTGAGTTA CTATCATTA TCAATGATAT GCCTATAACA AATGATCAGA AAAAATTAAT 6480  
GTCAAGCAAT GTTCAGATAG TAAGGCAACA AAGTTATTCC ATCATGTCTA TAATAAGGA 6540  
AGAAGTCCTT GCATATGTTG TACAGCTGCC TATCTATGGT GTAATAGATA CACCTTGCTG 6600  
GAAATTGCAC ACATCGCCTC TATGCACTAC CAACATCAAA GAAGGATCAA ATATTTGTTT 6660  
AACAAGGACT GATAGAGGAT GGTATTGTGA TAATGCAGGA TCAGTATCCT TCTTTCCACA 6720  
GGCTGACACT TGTAAAGTAC AGTCCAATCG AGTATTTTGT GACACTATGA ACAGTTTGAC 6780  
ATTACCAAGT GAAGTCAGCC TTTGTAACAC TGACATATTC AATTCCAAGT ATGACTGCAA 6840  
AATTATGACA TCAAAAACAG ACATAAGCAG CTCAGTAATT ACTTCTCTTG GAGCTATAGT 6900  
GTCATGCTAT GGTAAACTA AATGCACTGC ATCCAACAAA AATCGTGGGA TTATAAGAC 6960  
ATTTTCTAAT GGTGTGACT ATGTGTCAA CAAAGGAGTA GATACTGTGT CAGTGGGCAA 7020  
CACTTTTATAC TATGTAAACA AGCTGGAAGG CAAGAACCTT TATGTAAAAG GGGAACCTAT 7080  
AATAAATTAC TATGACCCTC TAGTGTTCCT TTCTGATGAG TTTGATGCAT CAATATCTCA 7140  
AGTCAATGAA AAAATCAATC AAAGTTTAGC TTTTATTCGT AGATCTGATG AATTACTACA 7200  
TAATGTAAAT ACTGGCAAAT CTAATAACAA TATTATGATA ACTACAATTA TTATAGTAAT 7260  
CATTGTAGTA TTGTTATCAT TAATAGCTAT TGGTTTACTG TTGTATTGTA AAGCCAAAAA 7320  
CACACCAGTT AACTAAGCA AAGACCAACT AAGTGGAATC AATAATATTG CATTACAGCA 7380  
ATAGACAAAA AACCACCTGA TCATGTTTCA ACAACAATCT GCTGACCACC AATCCCAAAT 7440

SUBSTITUTE SHEET (RULE 26)

- 388 -

CAACTTACAA CAAATATTTC AACATCACAG TACAGGCTGA ATCATTTCCT CACATCATGC 7500  
TACCCACATA ACTAAGCTAG ATCCTTAACT TATAGTTACA TAAAAACCTC AAGTATCACA 7560  
ATCAACCACT AAATCAACAC ATCATTACACA AAATTAACAG CTGGGGCAAA TATGTCGCGA 7620  
AGAAATCCTT GTAAATTGA GATTAGAGGT CATTGCTTGA ATGGTAGAAG ATGTCACTAC 7680  
AGTCATAATT ACTTTGAATG GCCTCCTCAT GCATTACTAG TGAGGCAAAA CTTCATGTTA 7740  
AACAAGATAC TCAAGTCAAT GGACAAAAGC ATAGACACTT TGTCTGAAAT AAGTGGAGCT 7800  
GCTGAACTGG ATAGAACAGA AGAATATGCT CTTGGTATAG TTGGAGTGCT AGAGAGTTAC 7860  
ATAGGATCTA TAAACAACAT AACAAAACAA TCAGCATGTG TTGCTATGAG TAACTTCTT 7920  
ATTGAGATCA ATAGTGATGA CATTAAAAAG CTTAGAGATA ATGAAGAACC CAATTCACCT 7980  
AAGATAAGAG TGTACAATAC TGTATATCA TACATTGAGA GCAATAGAAA AAACAACAAG 8040  
CAAACCATCC ATCTGCTCAA GAGACTACCA GCAGACGTGC TGAAGAAGAC AATAAAGAAC 8100  
ACATTAGATA TCCACAAAAG CATAACCATA AGCAATCCAA AAGAGTCAAC TGTGAATGAT 8160  
CAAAATGACC AAACCAAAAA TAATGATATT ACCGGATAAA TATCCTTGTA GTATATCATC 8220  
CATATTGATC TCAAGTGAAA GCATGGTTGC TACATTCAAT CATAAAAAACA TATTACAATT 8280  
TAACCATAAC TATTTGGATA ACCACCAGCG TTTATTAAAT CATATATTTG ATGAAATTCA 8340  
TTGGACACCT AAAAATTAT TAGATGCCAC TCAACAATTT CTCCAACATC TTAACATCCC 8400  
TGAAGATATA TATACAGTAT ATATATTAGT GTCATAATGC TTGACCATAA CGACTCTATG 8460  
TCATCCAACC ATAAACTAT TTTGATAAGG TTATGGGACA AAATGGATCC CATTATTAAT 8520  
GGAAACTCTG CTAATGTGTA TCTAAGTAT AGTTATTTAA AAGGTGTTAT CTCTTTTCA 8580  
GAGTGTAATG CTTTAGGGAG TTATCTTTTT AACGGCCCTT ATCTTAAAAA TGATTACACC 8640  
AACTTAATTA GTAGACAAAG CCCACTACTA GAGCATATGA ATCTTAAAAA ACTAACTATA 8700  
ACACAGTCAT TAATATCTAG ATATCATAAA GGTGAACTGA AATTAGAAGA ACCAACTTAT 8760  
TTCCAGTCAT TACTTATGAC ATATAAAGT ATGTCCTCGT CTGAACAAAT TGCTACAACT 8820  
AACTTACTTA AAAAAATAAT ACGAAGAGCC ATAGAAATAA GTGATGTAAA GGTGTACGCC 8880  
ATCTTGAATA AACTAGGATT AAAGGAAAAG GACAGAGTTA AGCCCAACAA TAATTCAGGT 8940  
GATGAAAAC CAGTACTTAC AACTATAATT AAAGATGATA TACTTTCGGC TGTGGAAAAC 9000

SUBSTITUTE SHEET (RULE 26)

- 389 -

AATCAATCAT ATACAAATTC AGACAAAAGT CACTCAGTAA ATCAAAATAT CACTATCAAA	9060
ACAACACTCT TGAAAAAATT GATGTGTTCA ATGCAACATC CTCCATCATG GTTAATACAC	9120
TGGTTCAATT TATATACAAA ATTAAATAAC ATATTAACAC AATATCGATC AAATGAGGTA	9180
AAAAGTCATG GGTTTATATT AATAGATAAT CAAACTTTAA GTGGTTTTCA GTTTATTTTA	9240
AATCAATATG GTTGATCGT TTATCATAAA GGA CTCAAAA AAATCACAAC TACTACTTAC	9300
AATCAATTTT TGACATGGAA AGACATCAGC CTTAGCAGAT TAAATGTTG CTTAATTACT	9360
TGGATAAGTA ATTGTTTTAA TACATTAAAC AAAAGCTTAG GGCTGAGATG TGGATTCAAT	9420
AATGTTGTGT TATACAATT ATTTCTTTAT GGAGATTGTA TACTGAAAT ATTTTCATAAT	9480
GAAGGCTTCT ACATAATAAA AGAAGTAGAG GGATTTATTA TGTCTTTAAT TCTAAACATA	9540
ACAGAAGAAG ATCAATTTAG GAAACGATT TATAATAGCA TGCTAAATAA CATCACAGAT	9600
GCAGCTATTA AGGCTCAAAA GGACCTACTA TCAAGAGTAT GTCACACTTT ATTAGACAAG	9660
ACAGTGTCTG ATAATATCAT AAATGGTAAA TGGATAATCC TATTAAGTAA ATTTCTTAAA	9720
TTGATTAAGC TTGCAGGTGA TAATAATCTC AATAACTTGA GTGAGCTATA TTTTCTCTTC	9780
AGAATCTTTG GACATCCAAT GGTGATGAA AGACAAGCAA TGGATTCTGT AAGAATTAAC	9840
TGTAATGAAA CTAAGTTCTA CTTATTAAGT AGTCTAAGTA CATTAAGAGG TGCTTTCATT	9900
TATAGAATCA TAAAAGGGTT TGTAATACC TACAACAGAT GGCCACCTT AAGGAATGCT	9960
ATTGTCCTAC CTCTAAGATG GTTAACTAC TATAACTTA ATACTTATCC ATCTCTACTT	10020
GAAATCACAG AAAATGATT GATTATTTTA TCAGGATTGC GGTTCTATCG TGAGTTTCAT	10080
CTGCCTAAAA AAGTGGATCT TGAAATGATA ATAAATGACA AAGCCATTC ACCTCCAAAA	10140
GATCTAATAT GGA CTAGTTT TCCTAGAAAT TACATGCCAT CACATATACA AAATTATATA	10200
GAACATGAAA AGTTGAAGTT CTCTGAAAGC GACAGATCGA GAAGAGTACT AGAGTATTAC	10260
TTGAGAGATA ATAAATTCAA TGAATGCGAT CTATACAATT GTGTAGTCAA TCAAAGCTAT	10320
CTCAACAACT CTAATCACGT GGTATCACTA ACTGGTAAAG AAAGAGAGCT CAGTGTAGGT	10380
AGAATGTTTG CTATGCAACC AGGTATGTTT AGGCAAATCC AAATCTTAGC AGAGAAAAATG	10440
ATAGCTGAAA ATATTTTACA ATTCTTCCCT GAGAGTTTGA CAAGATATGG TGATCTAGAG	10500
CTTCAAAAGA TATTAGAATT AAAAGCAGGA ATAAGCAACA AGTCAAATCG TTATAATGAT	10560

SUBSTITUTE SHEET (RULE 26)

- 390 -

AACTACAACA ATTATATCAG TAAATGTTCT ATCATTACAG ATCTTAGCAA ATTCAATCAG 10620  
GCATTTAGAT ATGAAACATC ATGTATCTGC AGTGATGTAT TAGATGAACT GCATGGAGTA 10680  
CAATCTCTGT TCTCTTGGTT GCATTAAACA ATACCTCTTG TCACAATAAT ATGTACATAT 10740  
AGACATGCAC CTCCTTTCAT AAAGGATCAT GTTGTTAATC TTAATGAGGT TGATGAACAA 10800  
AGTGGATTAT ACAGATATCA TATGGGTGGT ATTGAGGGCT GGTGTCAAAA ACTGTGGACC 10860  
ATTGAAGCTA TATCATTATT AGATCTAATA TCTCTCAAAG GGAAATTCTC TATCACAGCT 10920  
CTGATAAATG GTGATAATCA GTCAATTGAT ATAAGCAAAC CAGTTAGACT TATAGAGGGT 10980  
CAGACCCATG CACAAGCAGA TTATTTGTTA GCATTAAATA GCCTTAAATT GTTATATAAA 11040  
GAGTATGCAG GTATAGGCCA TAAGCTTAAG GGAACAGAGA CCTATATATC CCGAGATATC 11100  
CAGTTCATGA GCAAAACAAT CCAGCACAAT GGAGTGACT ATCCAGCCAG TATCAAAAAA 11160  
GTCCTGAGAG TAGGTCCATG GATAAACACG ATACTTGATG ATTTTAAAGT TAGTTTAGAA 11220  
TCTATAGGCA GCTTAACACA GGAGTTAGAA TACAGAGGAG AAAGCTTATT ATGCAGTTTA 11280  
ATATTTAGGA ACATTTGGTT ATACAATCAA ATTGCTTTGC AACTCCGAAA TCATGCATTA 11340  
TGTAACAATA AGCTATATTT AGATATATTG AAAGTATTAA AACACTTAAA AACTTTTTTT 11400  
AATCTTGATA GCATTGATAT GGCTTTATCA TTGTATATGA ATTTGCCTAT GCTGTTTGGT 11460  
GGTGGTGATC CTAATTTGTT ATATCGAAGC TTTTATAGGA GAACTCCAGA CTTCTTACA 11520  
GAAGCTATAG TACATTCAGT GTTTGTGTG AGCTATTATA CTGGTCACGA TTTACAAGAT 11580  
AAGCTCCAGG ATCTTCCAGA TGATAGACTG AACAAATTCT TGACATGTGT CATCACATTT 11640  
GATAAAAATC CCAATGCCGA GTTTGTAACA TTGATGAGGG ATCCACAGGC TTTAGGGTCT 11700  
GAAAGGCAAG CTAAAATTAC TAGTGAGATT AATAGATTAG CAGTAACAGA AGTCTTAAGT 11760  
ATAGCCCCAA ACAAATATT TTCTAAAAGT GCACAACATT ATACTACCAC TGAGATTGAT 11820  
CTAAATGACA TTATGCAAAA TATAGAACCA ACTTACCCTC ATGGATTAAG AGTTGTTTAT 11880  
GAAAGTTTAC CTTTTTATAA AGCAGAAAAA ATAGTTAATC TTATATCAGG AACAAAATCC 11940  
ATAACTAATA TACTTGAAAA AACATCAGCA ATAGATACAA CTGATATTAA TAGGGCTACT 12000  
GATATGATGA GGAAAAATAT AACTTTACTT ATAAGGATAC TTCCACTAGA TTGTAACAAA 12060  
GACAAAAGAG AGTTATTAAG TTTAGAAAAT CTTAGTATAA CTGAATTAAG CAAGTATGTA 12120

SUBSTITUTE SHEET (RULE 26)

- 391 -

AGAGAAAGAT CTTGGTCATT ATCCAATATA GTAGGAGTAA CATCGCCAAG TATTATGTTC 12180  
ACAATGGACA TTAAATATAC AACTAGCACT ATAGCCAGTG GTATAATAAT AGAAAAATAT 12240  
AATGTTAATA GTTTAACTCG TGGTGAAAGA GGACCCACCA AGCCATGGGT AGGCTCATCC 12300  
ACGCAGGAGA AAAAAACAAT GCCAGTGTAC AACAGACAAG TTTTAACCAA AAAGCAAAGA 12360  
GACCAAATAG ATTTATTAGC AAAATTAGAC TGGGTATATG CATCCATAGA CAACAAAGAT 12420  
GAATTCATGG AAGAACTGAG TACTGGAACA CTTGGACTGT CATATGAAAA AGCCAAAAAG 12480  
TTGTTTCCAC AATATCTAAG TGTCAATTAT TTACACCGTT TAACAGTCAG TAGTAGACCA 12540  
TGTGAATTCC CTGCATCAAT ACCAGCTTAT AGAACAAACA ATTATCATTT TGATACTAGT 12600  
CCTATCAATC ATGTATTAAC AGAAAAGTAT GGAGATGAAG ATATCGACAT TGTGTTTCAA 12660  
AATTGCATAA GTTTTGGTCT TAGCCTGATG TCGGTTGTGG AACAAATCAC AAACATATGT 12720  
CCTAATAGAA TTATTCTCAT ACCGAAGCTG AATGAGATAC ATTTGATGAA ACCTCCTATA 12780  
TTTACAGGAG ATGTTGATAT CATCAAGTTG AAGCAAGTGA TACAAAAGCA GCACATGTTC 12840  
CTACCAGATA AAATAAGTTT AACCCAATAT GTAGAATTAT TCTTAAGTAA CAAAGCACTT 12900  
AAATCTGGAT CTCACATCAA CTCTAATTTA ATATTAGTAC ATAAAATGTC TGATTATTTT 12960  
CATAATGCTT ATATTTTAAG TACTAATTTA GCTGGACATT GGATTCTGAT TATTCAACTT 13020  
ATGAAAGATT CAAAAGGTAT TTTTGAAAAA GATTGGGGAG AGGGGTACAT AACTGATCAT 13080  
ATGTTCATTA ATTTGAATGT TTTCTTTAAT GCTTATAAGA CTTATTTGCT ATGTTTTTCAT 13140  
AAAGGTTATG GTAAAGCAAA ATTAGAATGT GATATGAACA CTTGAGATCT TCTTTGTGTT 13200  
TTGGAGTTAA TAGACAGTAG CTAAGGAAA TCTATGTCTA AAGTTTTCCT AGAACAAAAA 13260  
GTCATAAAAT ACATAGTCAA TCAAGACACA AGTTTGCCTA GAATAAAAGG CTGTCACAGT 13320  
TTTAAGTTGT GGTTTTAAA ACGCCTTGAT AATGCTAAAT TTACCGTATG CCCTTGGGTT 13380  
GTTAACATAG ATTATCACCC AACACACATG AAAGCTATAT TATCTTACAT AGATTTAGTT 13440  
AGAATGGGGT TAATAAATGT AGATAAATTA ACCATTAAAA ATAAAAACAA ATTCAATGAT 13500  
GAATTTTACA CATCAAATCT CTTTACATT AGTTATAACT TTTCAGACAA CACTCATTTG 13560  
CTAACAAAAC AAATAAGAAT TGCTAATTCA GAATTAGAAG ATAATTATAA CAAACTATAT 13620  
CACCCAACCC CAGAACTTT AGAAAATATG TCATTAATTC CTGTTAAAAG TAATAATAGT 13680

SUBSTITUTE SHEET (RULE 26)

- 392 -

AACAAACCTA AATTTTGTAT AAGTGGAAAT ACCGAATCTA TGATGATGTC AACATTCTCT	13740
AGTAAAATGC ATATTAAATC TTCCACTGTT ACCACAAGAT TCAATTATAG CAAACAAGAC	13800
TTGTACAATT TATTTCCAAT TGTGTGATA GACAAGATTA TAGATCATTG AGGTAATACA	13860
GCAAAATCTA ACCAACTTTA CACCACCACT TCACATCAGA CATCTTTAGT AAGGAATAGT	13920
GCATCACTTT ATTGCATGCT TCCTTGCCAT CATGTCAATA GATTAACTT TGTATTTAGT	13980
TCCACAGGAT GCAAGATCAG TATAGAGTAT ATTTTAAAAG ATCTTAAGAT TAAGGACCCC	14040
AGTTGTATAG CATTATAGG TGAAGGAGCT GGTAACCTAT TATTACGTAC GGTAGTAGAA	14100
CTTCATCCAG ACATAAGATA CATTTACAGA AGTTTAAAAG ATTGCAATGA TCATAGTTTA	14160
CCTATTGAAT TTCTAAGGTT ATACAACGGG CATATAAACA TAGATTATGG TGAGAATTTA	14220
ACCATTCTCG CTACAGATGC AACTAATAAC ATTCATTGGT CTTATTTACA TATAAAATTT	14280
GCAGAACCTA TTAGCATCTT TGTCTGCGAT GCTGAATTAC CTGTTACAGC CAATTGGAGT	14340
AAAATTATAA TTGAATGGAG TAAGCATGTA AGAAAGTGCA AGTACTGTTT TTCTGTAAAT	14400
AGATGCATTT TAATTGCAAA ATATCATGCT CAAGATGACA TTGATTTCAA ATTAGATAAC	14460
ATTACTATAT TAAAACTTA CGTGTGCCA GGTAGCAAGT TAAAAGGATC TGAAGTTTAC	14520
TTAATCCTTA CAATAGGCCC TGCAAATATA CTTCTGTTT TTGATGTTGT ACAAATGCT	14580
AAATTGATAC TTTCAAGAAC TAAAAATTTT ATTATGCCTA AAAAACTGA CAAGGAATCT	14640
ATCGATGCAG TTATTAAAAG CTTAATACCT TTCCTTTGTT ACCCTATAAC AAAAAAGGA	14700
ATTAAGACTT CATTGTCAA ATTGAAGAGT GTAGTTAATG GAGATATATT ATCATATTCT	14760
ATAGCTGGAC GTAATGAAGT ATTCAGCAAC AAGCTTATAA ACCACAAGCA TATGAATATC	14820
CTAAAATGGC TAGATCATGT TTTAAATTTT AGATCAGCTG AACTTAATTA CAATCATTTA	14880
TACATGATAG AGTCCACATA TCCTTACTTA AGTGAATTGT TAAATAGTTT AACAACCAAT	14940
GAGCTCAAGA AGCTGATTAA AATAACAGGT AGTGTGCTAT ACAACCTTCC CAACGAACAG	15000
TAGTTTAAAA TATCATTAAC AAGTTTGGTC AAATTTAGAT GCTAACACAT CATTATATTA	15060
TAGTTATTAA AAAATATACA AACTTTTCAA TAATTTAGCA TATTGATTCC AAAATTATCA	15120
TTTAGTCTT AAGGGGTTAA ATAAAAGTCT AAAACTAACA ATTATACATG TGCATTCACA	15180
ACACAACGAG ACATTAGTTT TTGACACTTT TTTTCTCGT	15219

SUBSTITUTE SHEET (RULE 26)



- 393 -

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

```

Met Asp Pro Ile Ile Asn Gly Asn Ser Ala Asn Val Tyr Leu Thr Asp
1           5           10           15

Ser Tyr Leu Lys Gly Val Ile Ser Phe Ser Glu Cys Asn Ala Leu Gly
          20           25           30

Ser Tyr Leu Phe Asn Gly Pro Tyr Leu Lys Asn Asp Tyr Thr Asn Leu
          35           40           45

Ile Ser Arg Gln Ser Pro Leu Leu Glu His Met Asn Leu Lys Lys Leu
          50           55           60

Thr Ile Thr Gln Ser Leu Ile Ser Arg Tyr His Lys Gly Glu Leu Lys
65           70           75           80

Leu Glu Glu Pro Thr Tyr Phe Gln Ser Leu Leu Met Thr Tyr Lys Ser
          85           90           95

Met Ser Ser Ser Glu Gln Ile Ala Thr Thr Asn Leu Leu Lys Lys Ile
          100          105          110

Ile Arg Arg Ala Ile Glu Ile Ser Asp Val Lys Val Tyr Ala Ile Leu
          115          120          125

Asn Lys Leu Gly Leu Lys Glu Lys Asp Arg Val Lys Pro Asn Asn Asn
          130          135          140

Ser Gly Asp Glu Asn Ser Val Leu Thr Thr Ile Ile Lys Asp Asp Ile
145          150          155          160

Leu Ser Ala Val Glu Asn Asn Gln Ser Tyr Thr Asn Ser Asp Lys Ser
          165          170          175

His Ser Val Asn Gln Asn Ile Thr Ile Lys Thr Thr Leu Leu Lys Lys
          180          185          190

```

SUBSTITUTE SHEET (RULE 26)

- 394 -

Leu Met Cys Ser Met Gln His Pro Pro Ser Trp Leu Ile His Trp Phe  
 195 200 205  
 Asn Leu Tyr Thr Lys Leu Asn Asn Ile Leu Thr Gln Tyr Arg Ser Asn  
 210 215 220  
 Glu Val Lys Ser His Gly Phe Ile Leu Ile Asp Asn Gln Thr Leu Ser  
 225 230 235 240  
 Gly Phe Gln Phe Ile Leu Asn Gln Tyr Gly Cys Ile Val Tyr His Lys  
 245 250 255  
 Gly Leu Lys Lys Ile Thr Thr Thr Thr Tyr Asn Gln Phe Leu Thr Trp  
 260 265 270  
 Lys Asp Ile Ser Leu Ser Arg Leu Asn Val Cys Leu Ile Thr Trp Ile  
 275 280 285  
 Ser Asn Cys Leu Asn Thr Leu Asn Lys Ser Leu Gly Leu Arg Cys Gly  
 290 295 300  
 Phe Asn Asn Val Val Leu Ser Gln Leu Phe Leu Tyr Gly Asp Cys Ile  
 305 310 315 320  
 Leu Lys Leu Phe His Asn Glu Gly Phe Tyr Ile Ile Lys Glu Val Glu  
 325 330 335  
 Gly Phe Ile Met Ser Leu Ile Leu Asn Ile Thr Glu Glu Asp Gln Phe  
 340 345 350  
 Arg Lys Arg Phe Tyr Asn Ser Met Leu Asn Asn Ile Thr Asp Ala Ala  
 355 360 365  
 Ile Lys Ala Gln Lys Asp Leu Leu Ser Arg Val Cys His Thr Leu Leu  
 370 375 380  
 Asp Lys Thr Val Ser Asp Asn Ile Ile Asn Gly Lys Trp Ile Ile Leu  
 385 390 395 400  
 Leu Ser Lys Phe Leu Lys Leu Ile Lys Leu Ala Gly Asp Asn Asn Leu  
 405 410 415  
 Asn Asn Leu Ser Glu Leu Tyr Phe Leu Phe Arg Ile Phe Gly His Pro  
 420 425 430  
 Met Val Asp Glu Arg Gln Ala Met Asp Ser Val Arg Ile Asn Cys Asn  
 435 440 445  
 Glu Thr Lys Phe Tyr Leu Leu Ser Ser Leu Ser Thr Leu Arg Gly Ala  
 450 455 460  
 Phe Ile Tyr Arg Ile Ile Lys Gly Phe Val Asn Thr Tyr Asn Arg Trp

SUBSTITUTE SHEET (RULE 26)

- 395 -

465	470	475	480
Pro Thr Leu Arg Asn Ala Ile Val Leu Pro Leu Arg Trp Leu Asn Tyr			
	485	490	495
Tyr Lys Leu Asn Thr Tyr Pro Ser Leu Leu Glu Ile Thr Glu Asn Asp			
	500	505	510
Leu Ile Ile Leu Ser Gly Leu Arg Phe Tyr Arg Glu Phe His Leu Pro			
	515	520	525
Lys Lys Val Asp Leu Glu Met Ile Ile Asn Asp Lys Ala Ile Ser Pro			
	530	535	540
Pro Lys Asp Leu Ile Trp Thr Ser Phe Pro Arg Asn Tyr Met Pro Ser			
	545	550	555
His Ile Gln Asn Tyr Ile Glu His Glu Lys Leu Lys Phe Ser Glu Ser			
	565	570	575
Asp Arg Ser Arg Arg Val Leu Glu Tyr Tyr Leu Arg Asp Asn Lys Phe			
	580	585	590
Asn Glu Cys Asp Leu Tyr Asn Cys Val Val Asn Gln Ser Tyr Leu Asn			
	595	600	605
Asn Ser Asn His Val Val Ser Leu Thr Gly Lys Glu Arg Glu Leu Ser			
	610	615	620
Val Gly Arg Met Phe Ala Met Gln Pro Gly Met Phe Arg Gln Ile Gln			
	625	630	635
Ile Leu Ala Glu Lys Met Ile Ala Glu Asn Ile Leu Gln Phe Phe Pro			
	645	650	655
Glu Ser Leu Thr Arg Tyr Gly Asp Leu Glu Leu Gln Lys Ile Leu Glu			
	660	665	670
Leu Lys Ala Gly Ile Ser Asn Lys Ser Asn Arg Tyr Asn Asp Asn Tyr			
	675	680	685
Asn Asn Tyr Ile Ser Lys Cys Ser Ile Ile Thr Asp Leu Ser Lys Phe			
	690	695	700
Asn Gln Ala Phe Arg Tyr Glu Thr Ser Cys Ile Cys Ser Asp Val Leu			
	705	710	715
Asp Glu Leu His Gly Val Gln Ser Leu Phe Ser Trp Leu His Leu Thr			
	725	730	735
Ile Pro Leu Val Thr Ile Ile Cys Thr Tyr Arg His Ala Pro Pro Phe			
	740	745	750

SUBSTITUTE SHEET (RULE 26)

- 396 -

Ile Lys Asp His Val Val Asn Leu Asn Glu Val Asp Glu Gln Ser Gly  
 755 760 765  
 Leu Tyr Arg Tyr His Met Gly Gly Ile Glu Gly Trp Cys Gln Lys Leu  
 770 775 780  
 Trp Thr Ile Glu Ala Ile Ser Leu Leu Asp Leu Ile Ser Leu Lys Gly  
 785 790 795 800  
 Lys Phe Ser Ile Thr Ala Leu Ile Asn Gly Asp Asn Gln Ser Ile Asp  
 805 810 815  
 Ile Ser Lys Pro Val Arg Leu Ile Glu Gly Gln Thr His Ala Gln Ala  
 820 825 830  
 Asp Tyr Leu Leu Ala Leu Asn Ser Leu Lys Leu Leu Tyr Lys Glu Tyr  
 835 840 845  
 Ala Gly Ile Gly His Lys Leu Lys Gly Thr Glu Thr Tyr Ile Ser Arg  
 850 855 860  
 Asp Met Gln Phe Met Ser Lys Thr Ile Gln His Asn Gly Val Tyr Tyr  
 865 870 875 880  
 Pro Ala Ser Ile Lys Lys Val Leu Arg Val Gly Pro Trp Ile Asn Thr  
 885 890 895  
 Ile Leu Asp Asp Phe Lys Val Ser Leu Glu Ser Ile Gly Ser Leu Thr  
 900 905 910  
 Gln Glu Leu Glu Tyr Arg Gly Glu Ser Leu Leu Cys Ser Leu Ile Phe  
 915 920 925  
 Arg Asn Ile Trp Leu Tyr Asn Gln Ile Ala Leu Gln Leu Arg Asn His  
 930 935 940  
 Ala Leu Cys Asn Asn Lys Leu Tyr Leu Asp Ile Leu Lys Val Leu Lys  
 945 950 955 960  
 His Leu Lys Thr Phe Phe Asn Leu Asp Ser Ile Asp Met Ala Leu Ser  
 965 970 975  
 Leu Tyr Met Asn Leu Pro Met Leu Phe Gly Gly Gly Asp Pro Asn Leu  
 980 985 990  
 Leu Tyr Arg Ser Phe Tyr Arg Arg Thr Pro Asp Phe Leu Thr Glu Ala  
 995 1000 1005  
 Ile Val His Ser Val Phe Val Leu Ser Tyr Tyr Thr Gly His Asp Leu  
 1010 1015 1020

SUBSTITUTE SHEET (RULE 26)

- 397 -

Gln Asp Lys Leu Gln Asp Leu Pro Asp Asp Arg Leu Asn Lys Phe Leu  
 1025 1030 1035 1040  
 Thr Cys Val Ile Thr Phe Asp Lys Asn Pro Asn Ala Glu Phe Val Thr  
 1045 1050 1055  
 Leu Met Arg Asp Pro Gln Ala Leu Gly Ser Glu Arg Gln Ala Lys Ile  
 1060 1065 1070  
 Thr Ser Glu Ile Asn Arg Leu Ala Val Thr Glu Val Leu Ser Ile Ala  
 1075 1080 1085  
 Pro Asn Lys Ile Phe Ser Lys Ser Ala Gln His Tyr Thr Thr Thr Glu  
 1090 1095 1100  
 Ile Asp Leu Asn Asp Ile Met Gln Asn Ile Glu Pro Thr Tyr Pro His  
 1105 1110 1115 1120  
 Gly Leu Arg Val Val Tyr Glu Ser Leu Pro Phe Tyr Lys Ala Glu Lys  
 1125 1130 1135  
 Ile Val Asn Leu Ile Ser Gly Thr Lys Ser Ile Thr Asn Ile Leu Glu  
 1140 1145 1150  
 Lys Thr Ser Ala Ile Asp Thr Thr Asp Ile Asn Arg Ala Thr Asp Met  
 1155 1160 1165  
 Met Arg Lys Asn Ile Thr Leu Leu Ile Arg Ile Leu Pro Leu Asp Cys  
 1170 1175 1180  
 Asn Lys Asp Lys Arg Glu Leu Leu Ser Leu Glu Asn Leu Ser Ile Thr  
 1185 1190 1195 1200  
 Glu Leu Ser Lys Tyr Val Arg Glu Arg Ser Trp Ser Leu Ser Asn Ile  
 1205 1210 1215  
 Val Gly Val Thr Ser Pro Ser Ile Met Phe Thr Met Asp Ile Lys Tyr  
 1220 1225 1230  
 Thr Thr Ser Thr Ile Ala Ser Gly Ile Ile Ile Glu Lys Tyr Asn Val  
 1235 1240 1245  
 Asn Ser Leu Thr Arg Gly Glu Arg Gly Pro Thr Lys Pro Trp Val Gly  
 1250 1255 1260  
 Ser Ser Thr Gln Glu Lys Lys Thr Met Pro Val Tyr Asn Arg Gln Val  
 1265 1270 1275 1280  
 Leu Thr Lys Lys Gln Arg Asp Gln Ile Asp Leu Leu Ala Lys Leu Asp  
 1285 1290 1295  
 Trp Val Tyr Ala Ser Ile Asp Asn Lys Asp Glu Phe Met Glu Glu Leu

SUBSTITUTE SHEET (RULE 26)

- 398 -

1300	1305	1310
Ser Thr Gly Thr Leu Gly Leu Ser Tyr Glu Lys Ala Lys Lys Leu Phe 1315	1320	1325
Pro Gln Tyr Leu Ser Val Asn Tyr Leu His Arg Leu Thr Val Ser Ser 1330	1335	1340
Arg Pro Cys Glu Phe Pro Ala Ser Ile Pro Ala Tyr Arg Thr Thr Asn 1345	1350	1355 1360
Tyr His Phe Asp Thr Ser Pro Ile Asn His Val Leu Thr Glu Lys Tyr 1365	1370	1375
Gly Asp Glu Asp Ile Asp Ile Val Phe Gln Asn Cys Ile Ser Phe Gly 1380	1385	1390
Leu Ser Leu Met Ser Val Val Glu Gln Phe Thr Asn Ile Cys Pro Asn 1395	1400	1405
Arg Ile Ile Leu Ile Pro Lys Leu Asn Glu Ile His Leu Met Lys Pro 1410	1415	1420
Pro Ile Phe Thr Gly Asp Val Asp Ile Ile Lys Leu Lys Gln Val Ile 1425	1430	1435 1440
Gln Lys Gln His Met Phe Leu Pro Asp Lys Ile Ser Leu Thr Gln Tyr 1445	1450	1455
Val Glu Leu Phe Leu Ser Asn Lys Ala Leu Lys Ser Gly Ser His Ile 1460	1465	1470
Asn Ser Asn Leu Ile Leu Val His Lys Met Ser Asp Tyr Phe His Asn 1475	1480	1485
Ala Tyr Ile Leu Ser Thr Asn Leu Ala Gly His Trp Ile Leu Ile Ile 1490	1495	1500
Gln Leu Met Lys Asp Ser Lys Gly Ile Phe Glu Lys Asp Trp Gly Glu 1505	1510	1515 1520
Gly Tyr Ile Thr Asp His Met Phe Ile Asn Leu Asn Val Phe Phe Asn 1525	1530	1535
Ala Tyr Lys Thr Tyr Leu Leu Cys Phe His Lys Gly Tyr Gly Lys Ala 1540	1545	1550
Lys Leu Glu Cys Asp Met Asn Thr Ser Asp Leu Leu Cys Val Leu Glu 1555	1560	1565
Leu Ile Asp Ser Ser Tyr Trp Lys Ser Met Ser Lys Val Phe Leu Glu 1570	1575	1580

SUBSTITUTE SHEET (RULE 26)

- 399 -

Gln Lys Val Ile Lys Tyr Ile Val Asn Gln Asp Thr Ser Leu Arg Arg  
 1585 1590 1595 1600  
 Ile Lys Gly Cys His Ser Phe Lys Leu Trp Phe Leu Lys Arg Leu Asp  
 1605 1610 1615  
 Asn Ala Lys Phe Thr Val Cys Pro Trp Val Val Asn Ile Asp Tyr His  
 1620 1625 1630  
 Pro Thr His Met Lys Ala Ile Leu Ser Tyr Ile Asp Leu Val Arg Met  
 1635 1640 1645  
 Gly Leu Ile Asn Val Asp Lys Leu Thr Ile Lys Asn Lys Asn Lys Phe  
 1650 1655 1660  
 Asn Asp Glu Phe Tyr Thr Ser Asn Leu Phe Tyr Ile Ser Tyr Asn Phe  
 1665 1670 1675 1680  
 Ser Asp Asn Thr His Leu Leu Thr Lys Gln Ile Arg Ile Ala Asn Ser  
 1685 1690 1695  
 Glu Leu Glu Asp Asn Tyr Asn Lys Leu Tyr His Pro Thr Pro Glu Thr  
 1700 1705 1710  
 Leu Glu Asn Met Ser Leu Ile Pro Val Lys Ser Asn Asn Ser Asn Lys  
 1715 1720 1725  
 Pro Lys Phe Cys Ile Ser Gly Asn Thr Glu Ser Met Met Met Ser Thr  
 1730 1735 1740  
 Phe Ser Ser Lys Met His Ile Lys Ser Ser Thr Val Thr Thr Arg Phe  
 1745 1750 1755 1760  
 Asn Tyr Ser Lys Gln Asp Leu Tyr Asn Leu Phe Pro Ile Val Val Ile  
 1765 1770 1775  
 Asp Lys Ile Ile Asp His Ser Gly Asn Thr Ala Lys Ser Asn Gln Leu  
 1780 1785 1790  
 Tyr Thr Thr Thr Ser His Gln Thr Ser Leu Val Arg Asn Ser Ala Ser  
 1795 1800 1805  
 Leu Tyr Cys Met Leu Pro Trp His His Val Asn Arg Phe Asn Phe Val  
 1810 1815 1820  
 Phe Ser Ser Thr Gly Cys Lys Ile Ser Ile Glu Tyr Ile Leu Lys Asp  
 1825 1830 1835 1840  
 Leu Lys Ile Lys Asp Pro Ser Cys Ile Ala Phe Ile Gly Glu Gly Ala  
 1845 1850 1855

SUBSTITUTE SHEET (RULE 26)

- 400 -

Gly Asn Leu Leu Leu Arg Thr Val Val Glu Leu His Pro Asp Ile Arg  
 1860 1865 1870  
 Tyr Ile Tyr Arg Ser Leu Lys Asp Cys Asn Asp His Ser Leu Pro Ile  
 1875 1880 1885  
 Glu Phe Leu Arg Leu Tyr Asn Gly His Ile Asn Ile Asp Tyr Gly Glu  
 1890 1895 1900  
 Asn Leu Thr Ile Pro Ala Thr Asp Ala Thr Asn Asn Ile His Trp Ser  
 1905 1910 1915 1920  
 Tyr Leu His Ile Lys Phe Ala Glu Pro Ile Ser Ile Phe Val Cys Asp  
 1925 1930 1935  
 Ala Glu Leu Pro Val Thr Ala Asn Trp Ser Lys Ile Ile Ile Glu Trp  
 1940 1945 1950  
 Ser Lys His Val Arg Lys Cys Lys Tyr Cys Ser Ser Val Asn Arg Cys  
 1955 1960 1965  
 Ile Leu Ile Ala Lys Tyr His Ala Gln Asp Asp Ile Asp Phe Lys Leu  
 1970 1975 1980  
 Asp Asn Ile Thr Ile Leu Lys Thr Tyr Val Cys Leu Gly Ser Lys Leu  
 1985 1990 1995 2000  
 Lys Gly Ser Glu Val Tyr Leu Ile Leu Thr Ile Gly Pro Ala Asn Ile  
 2005 2010 2015  
 Leu Pro Val Phe Asp Val Val Gln Asn Ala Lys Leu Ile Leu Ser Arg  
 2020 2025 2030  
 Thr Lys Asn Phe Ile Met Pro Lys Lys Thr Asp Lys Glu Ser Ile Asp  
 2035 2040 2045  
 Ala Val Ile Lys Ser Leu Ile Pro Phe Leu Cys Tyr Pro Ile Thr Lys  
 2050 2055 2060  
 Lys Gly Ile Lys Thr Ser Leu Ser Lys Leu Lys Ser Val Val Asn Gly  
 2065 2070 2075 2080  
 Asp Ile Leu Ser Tyr Ser Ile Ala Gly Arg Asn Glu Val Phe Ser Asn  
 2085 2090 2095  
 Lys Leu Ile Asn His Lys His Met Asn Ile Leu Lys Trp Leu Asp His  
 2100 2105 2110  
 Val Leu Asn Phe Arg Ser Ala Glu Leu Asn Tyr Asn His Leu Tyr Met  
 2115 2120 2125  
 Ile Glu Ser Thr Tyr Pro Tyr Leu Ser Glu Leu Leu Asn Ser Leu Thr

SUBSTITUTE SHEET (RULE 26)



- 401 -

2130	2135	2140
Thr Asn Glu Leu Lys Lys Leu Ile Lys Ile Thr Gly Ser Val Leu Tyr		
2145	2150	2155 2160
Asn Leu Pro Asn Glu Gln		
2165		

## (2) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CATATCACTC ACTCTGGGAT GGAG

24

## (2) INFORMATION FOR SEQ ID NO:36:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

TCAGAACATC AAGCACCGCC

20

## (2) INFORMATION FOR SEQ ID NO:37:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

SUBSTITUTE SHEET (RULE 26)

- 402 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ACAGTCAAGA CTGAGATGAG

20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AAGAGTCAGA TACATGTGGA

20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ACATGAATCA GCCTAAAGTC

20

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

SUBSTITUTE SHEET (RULE 26)

- 403 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CCGAAAGAGT TCCTGCGTTA CGACC

25

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

CAGTCCACAC AAGTACCAGG

20

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GTCAGAAGCT GTGGACCATC

20

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

- 404 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AATATTGCTA CAACAATGGC

20

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

ACTCTTCATT CCTAGACTGG

20

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

GTCCAATTAT GACTATGAAC

20

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

- 405 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AGAACAGACA TGAAGCTTGC

20

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CCAACAAGGA ATGCTTCTAG

20

(2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

ACAGCACTAT CTATGATTGA CCTGG

25

(2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

SUBSTITUTE SHEET (RULE 26)

- 406 -

GCAACATGGT TTACACATGC

20

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGATTGAGAG TTGATCCAGG

20

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AGGAGATACT TAAACTAAGC

20

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

SUBSTITUTE SHEET (RULE 26)

- 407 -

TAAGCTTATG CCTTTCAGCG

20

## (2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TTAACGGACC TAAGCTGTGC

20

## (2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GAAACAGATT ATTATGACGG

20

## (2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CGGGCTATCT AGGTGAACTT CAGG

24

SUBSTITUTE SHEET (RULE 26)

- 408 -

## (2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

ATTGGATAT GGAATATGAG

20

## (2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

ACTCAACTGA ACTACCAGTG

20

## (2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AAGAACATCA TGTATTTCAG

20



- 409 -

## (2) INFORMATION FOR SEQ ID NO:59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTATCAACGC ACTGCTCATG

20

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ATTTTCAGCA ATCACTTGGC ATGCC

25

## (2) INFORMATION FOR SEQ ID NO:61:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GCCTCTGTGC AAACAAGCTG

20

## (2) INFORMATION FOR SEQ ID NO:62:

SUBSTITUTE SHEET (RULE 26)

- 410 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TCTCTAGTTA CTCTAGCAGC

20

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

AGGTCGTTGT TTGTGAGGAG

20

## (2) INFORMATION FOR SEQ ID NO:64:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TCGTCCTCTT CTTTACTGTC

20

## (2) INFORMATION FOR SEQ ID NO:65:

SUBSTITUTE SHEET (RULE 26)

- 411 -

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CCGTCCTCGA GCTAGCCTCG

20

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CTCCTCCAGG CTCACATTGG

20

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GGGTTGGTAC ATAGCTCTGC

20

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:

SUBSTITUTE SHEET (RULE 26)

- 412 -

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CACCCATCTG ATATTTCCCT GATGG

25

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TGTTGACAG TACAAATCTG

20

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CTGAAATGGG AAGATTGTGC

20

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs

SUBSTITUTE SHEET (RULE 26)

- 413 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

AGCAATCTAC ACTGCCTACC

20

(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

TCACAGATGA TTCAATTATC

20

(2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GATCCTAGAT ATAAGTTCTC

20

(2) INFORMATION FOR SEQ ID NO:74:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid

SUBSTITUTE SHEET (RULE 26)

- 414 -

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

ACCAAACAAA GTTGGGTAAG G

21

(2) INFORMATION FOR SEQ ID NO:75:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GGGGGATCCA TCCCTAATCC TGCTCTTGTC CC

32

(2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GATTCCTCTG ATGGCTCCAC

20

(2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

SUBSTITUTE SHEET (RULE 26)

- 415 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TAACAGTCAA GGAGACCAAA G

21

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

GGGAAGCTTA ACCCTAATCC TGCCCTAGGT GG

32

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

ACCAGACAAA GCTGGGAATA GA

22

- 416 -

What is claimed is:

1. An isolated, recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene.
2. The virus of Claim 1 wherein the virus is from the Family Paramyxoviridae.
3. The virus of Claim 2 wherein the virus is from the Subfamily Paramyxovirinae.
4. The virus of Claim 3 wherein the virus is from the Genus Morbillivirus.
5. The virus of Claim 4 wherein the virus is measles virus.
6. The measles virus of Claim 5 wherein:
  - (a) the at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 26 (A → T), nucleotide 42 (A → T or A → C) and nucleotide 96 (G → A), where these nucleotides are presented in positive strand, antigenomic, message sense; and
  - (b) the at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 331 (isoleucine → threonine), 1409 (alanine → threonine), 1624 (threonine → alanine), 1649 (arginine → methionine), 1717



- 417 -

(aspartic acid → alanine), 1936  
(histidine → tyrosine), 2074  
(glutamine → arginine) and 2114  
(arginine → lysine).

7. The virus of Claim 3 wherein the virus is from the Genus Paramyxovirus.

8. The virus of Claim 7 wherein the virus is human parainfluenzae virus type 3 (PIV-3).

9. The PIV-3 of Claim 8 wherein:

- (a) the at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 23 (T → C), nucleotide 24 (C → T), nucleotide 28 (G → T) and nucleotide 45 (T → A), where these nucleotides are presented in positive strand, antigenomic, message sense; and
- (b) the at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 942 (tyrosine → histidine), 992 (leucine → phenylalanine), 1292 (leucine → phenylalanine), and 1558 (threonine → isoleucine).

10. The virus of Claim 3 wherein the virus is from the Genus Rubulavirus.

11. The virus of Claim 2 wherein the virus is from the Subfamily Pneumovirinae.

12. The virus of Claim 11 wherein the virus is from the Genus Pneumovirus.

SUBSTITUTE SHEET (RULE 26)

- 418 -

13. The virus of Claim 12 wherein the virus is human respiratory syncytial virus (RSV) subgroup B.

14. The virus of Claim 13 wherein:

- (a) the at least one attenuating mutation in the 3' genomic promoter region is selected from the group consisting of nucleotide 4 (C → G) and the insertion of an additional A in the stretch of A's at nucleotides 6-11, where these nucleotides are presented in positive strand, antigenomic, message sense; and
- (b) the at least one attenuating mutation in the RNA polymerase gene is selected from the group consisting of nucleotide changes which produce changes in an amino acid selected from the group consisting of residues 353 (arginine → lysine), 451 (lysine → arginine), 1229 (aspartic acid → asparagine), 2029 (threonine → isoleucine) and 2050 (asparagine → aspartic acid).

15. The virus of Claim 1 wherein the virus is from the Family Rhabdoviridae.

16. The virus of Claim 1 wherein the virus is from the Family Filoviridae.

17. A vaccine comprising an isolated, recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales according to Claim 1 and a physiologically acceptable carrier.

18. The vaccine of Claim 17 comprising a measles virus according to Claim 5 and a physiologically acceptable carrier.

SUBSTITUTE SHEET (RULE 26)

- 419 -

19. The vaccine of Claim 18 comprising a measles virus according to Claim 6 and a physiologically acceptable carrier.

20. The vaccine of Claim 17 comprising a PIV-3 according to Claim 8 and a physiologically acceptable carrier.

21. The vaccine of Claim 20 comprising a PIV-3 according to Claim 9 and a physiologically acceptable carrier.

22. The vaccine of Claim 17 comprising an RSV subgroup B according to Claim 13 and a physiologically acceptable carrier.

23. The vaccine of Claim 22 comprising an RSV subgroup B according to Claim 14 and a physiologically acceptable carrier.

24. A method for immunizing an individual to induce protection against a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales which comprises administering to the individual the vaccine of Claim 17.

25. The method of Claim 24 wherein the vaccine is the vaccine of Claim 18.

26. The method of Claim 25 wherein the vaccine is the vaccine of Claim 19.

27. The method of Claim 24 wherein the vaccine is the vaccine of Claim 20.

28. The method of Claim 27 wherein the vaccine is the vaccine of Claim 21.

29. The method of Claim 24 wherein the vaccine is the vaccine of Claim 22.

30. The method of Claim 29 wherein the vaccine is the vaccine of Claim 23.

- 420 -

31. An isolated nucleic acid molecule comprising a measles virus sequence in positive strand, antigenomic message sense selected from the group consisting of 1977 wild-type strain (SEQ ID NO:3), 1983 wild-type strain (SEQ ID NO:5) where the nucleotide 2499 is G or C, Montefiore wild-type strain (SEQ ID NO:7), Rubeovax™ vaccine strain (SEQ ID NO:9), where the nucleotide 2143 is T or C, Moraten vaccine strain (SEQ ID NO:11), Schwarz vaccine strain (SEQ ID NO:11), where the nucleotide 4917 is C and the nucleotide 4924 is C, and Zagreb vaccine strain (SEQ ID NO:13), and the complementary genomic sequences thereof.

32. An isolated nucleic acid molecule comprising a PIV-3 sequence in positive strand, antigenomic message sense selected from the group consisting of cp45 vaccine strain grown in fetal rhesus lung cells (SEQ ID NO:19) and cp45 vaccine strain grown in Vero cells (SEQ ID NO:21), and the complementary genomic sequences thereof.

33. A composition which comprises a transcription vector comprising an isolated nucleic acid molecule encoding a genome or antigenome of a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene, together with at least one expression vector which comprises at least one isolated nucleic acid molecule encoding the trans-acting proteins necessary for encapsidation, transcription and replication, whereby upon expression an infectious attenuated virus is produced.

- 421 -

34. The composition of Claim 33 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes a measles virus according to Claim 5 and the at least one expression vector comprises at least one isolated nucleic acid molecule encoding the trans-acting proteins N, P and L.

35. The composition of Claim 34 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes a measles virus according to Claim 6.

36. The composition of Claim 33 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes a PIV-3 according to Claim 8 and the at least one expression vector comprises at least one isolated nucleic acid molecule encoding the trans-acting proteins NP, P and L.

37. The composition of Claim 36 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes a PIV-3 according to Claim 9.

38. The composition of Claim 33 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes an RSV subgroup B according to Claim 13 and the at least one expression vector comprises at least one isolated nucleic acid molecule encoding the trans-acting proteins N, P, L and M2.

39. The composition of Claim 38 wherein the transcription vector comprises an isolated nucleic acid molecule which encodes an RSV subgroup B according to Claim 14.

40. A method for producing infectious attenuated nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales which comprises transforming or transfecting host cells with

- 422 -

the at least two vectors of Claim 33 and culturing the host cells under conditions which permit the co-expression of these vectors so as to produce the infectious attenuated virus.

41. The method of Claim 40 wherein the virus is the measles virus of Claim 5.

42. The method of Claim 41 wherein the virus is the measles virus of Claim 6.

43. The method of Claim 40 wherein the virus is the PIV-3 of Claim 8.

44. The method of Claim 43 wherein the virus is the PIV-3 of Claim 9.

45. The method of Claim 40 wherein the virus is the RSV subgroup B of Claim 13.

46. The method of Claim 45 wherein the virus is the RSV subgroup B of Claim 14.

1/3

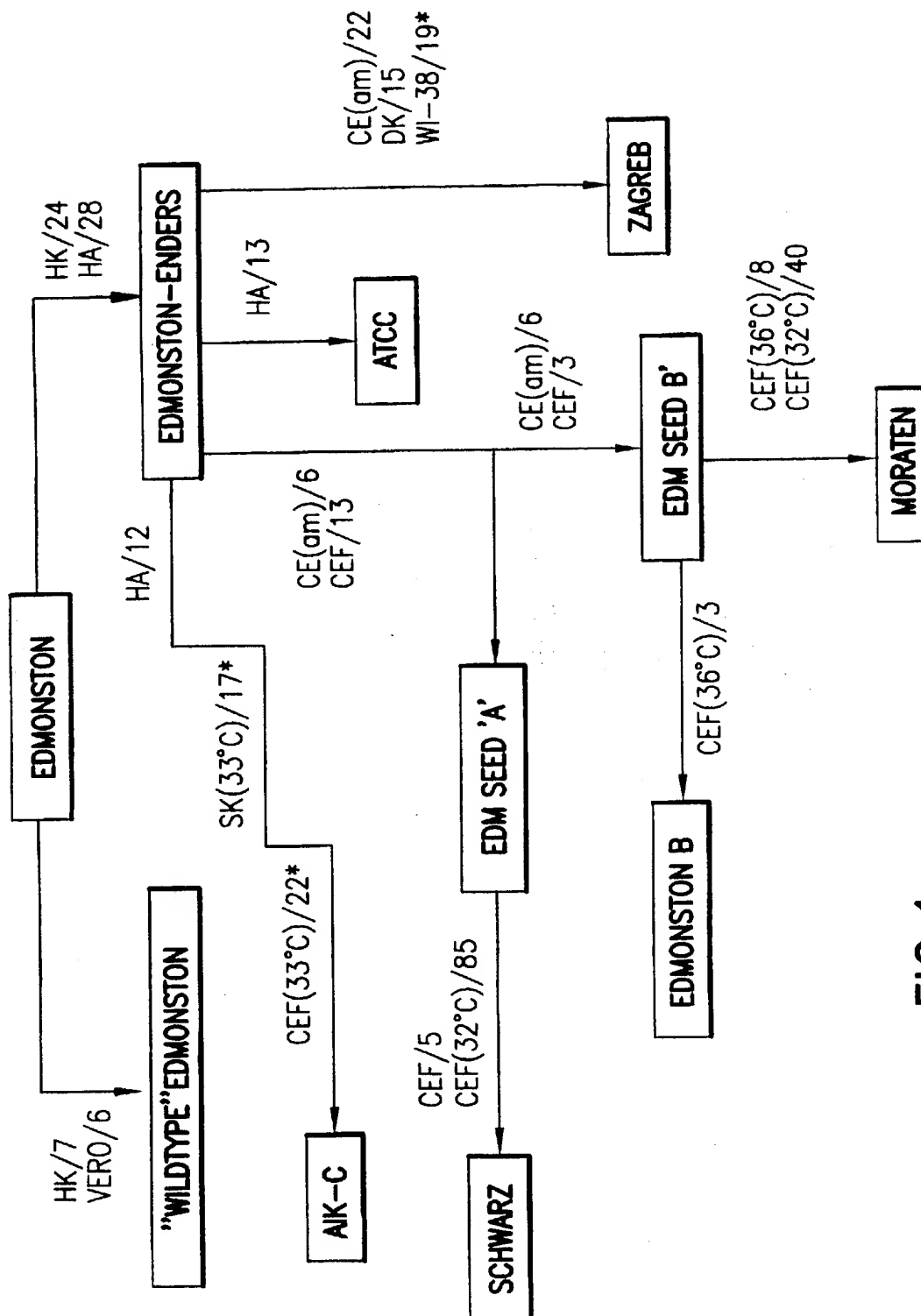


FIG. 1

2/3



PUTATIVE EXTENDED PROMOTER AND REGULATORY REGIONS OF GENOME AND ANTIGENOME:  
HIGHLY CONSERVED CIS-ACTING DOMAINS

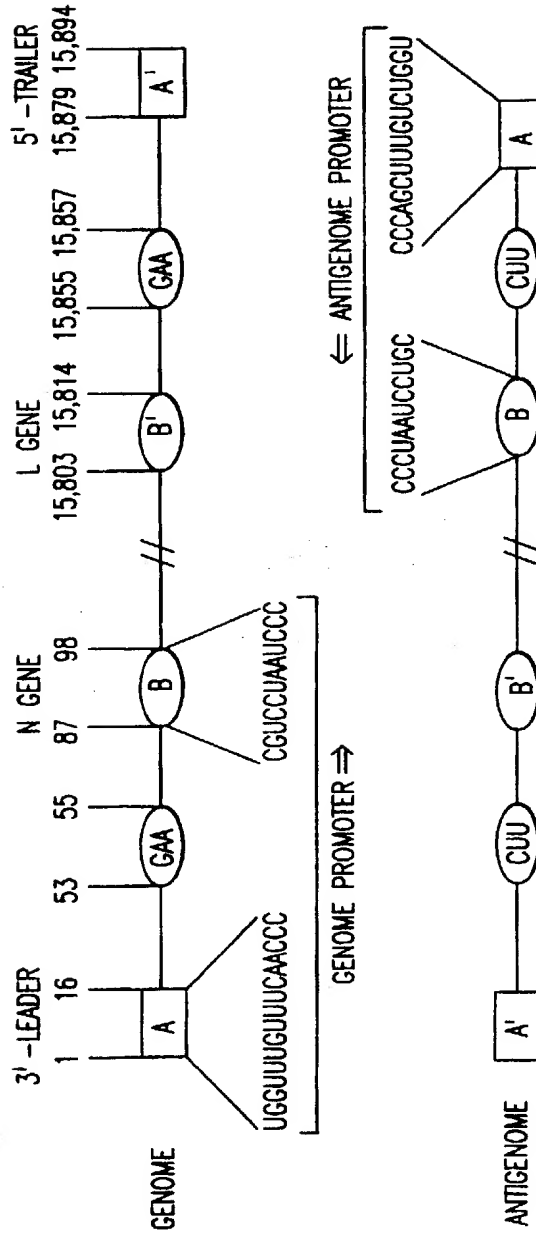


FIG.2



3/3

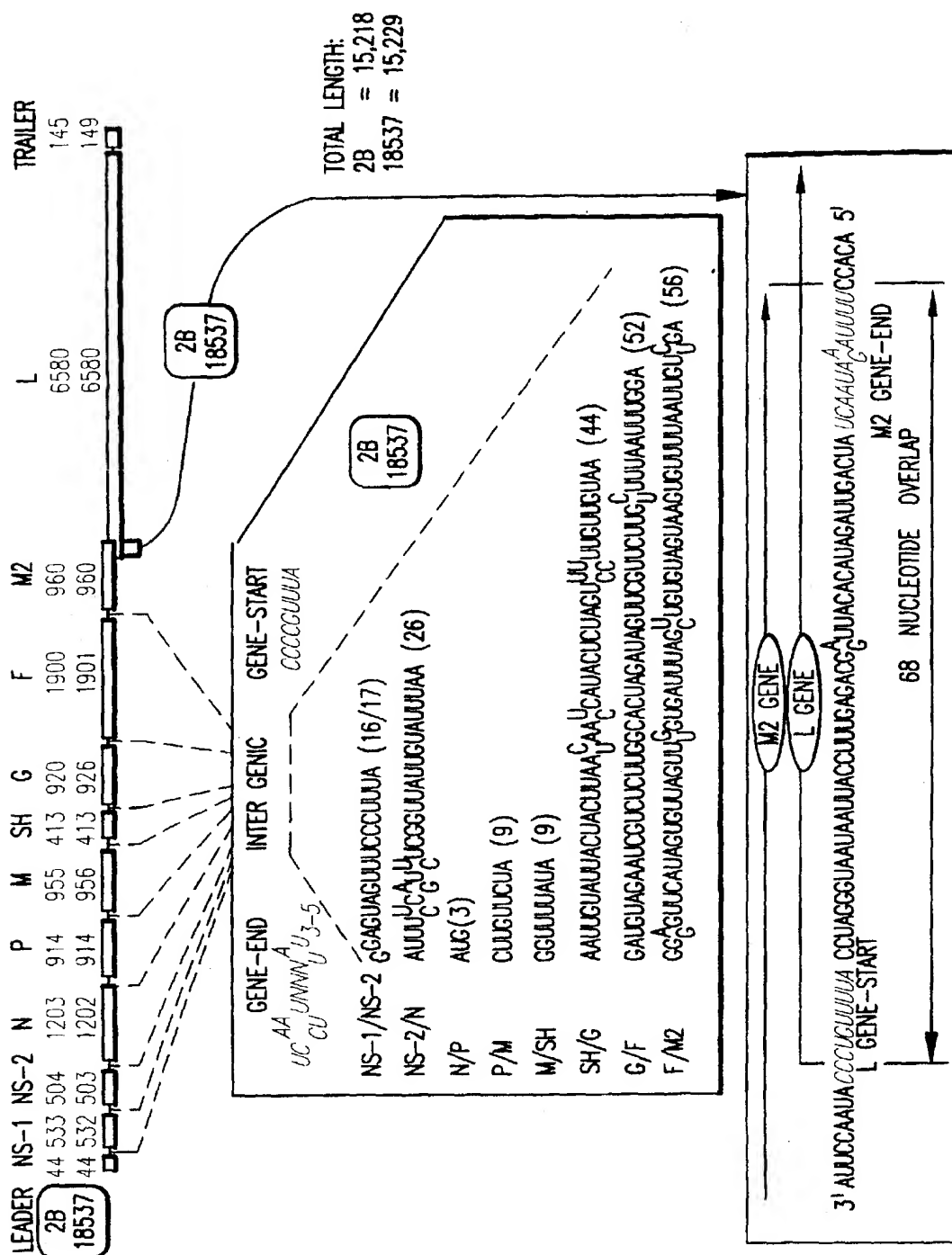


FIG.3

SUBSTITUTE SHEET (RULE 26)





## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification <sup>6</sup> : C12N 15/45, 15/47, 15/40, C07K 14/12, 14/115, 14/135, C12N 7/04, A61K 39/165, 39/155, C07K 14/145, 14/08</p>	<p><b>A3</b></p>	<p>(11) International Publication Number: <b>WO 98/13501</b></p> <p>(43) International Publication Date: 2 April 1998 (02.04.98)</p>
<p>(21) International Application Number: PCT/US97/16718</p> <p>(22) International Filing Date: 19 September 1997 (19.09.97)</p> <p>(30) Priority Data: 60/026,823 27 September 1996 (27.09.96) US</p> <p>(71) Applicants (for all designated States except US): AMERICAN CYANAMID COMPANY [US/US]; Five Giralda Farms, Madison, NJ 07940 (US). THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by THE DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Suite 325, 6011 Executive Boulevard, Rockville, MD 20852 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): UDEM, Stephen, A. [US/US]; Apartment 6F/G, 155 West 70th Street, New York, NY 10023 (US). SIDHU, Mohinderjit, S. [US/US]; 3 Aspen Court, Highland Mills, NY 10930 (US). TATEM, Joanne, M. [US/US]; 62 Douglas Drive, Towaco, NJ 07082 (US). MURPHY, Brian, R. [US/US]; 5410 Tuscarawas Road, Bethesda, MD 20816 (US). RANDOLPH, Valerie, B. [US/US]; 535 Pine Brook Road, Lincoln Park, NJ 07035 (US).</p>	<p>(74) Agents: GORDON, Alan, M. et al.; American Home Products Corporation, Patent Law Dept. - 2B, One Campus Drive, Parsippany, NJ 07054 (US).</p> <p>(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, RO, RU, SD, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ZW, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p><b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p> <p>(88) Date of publication of the international search report: 13 August 1998 (13.08.98)</p>	
<p>(54) Title: 3' GENOMIC PROMOTER REGION AND POLYMERASE GENE MUTATIONS RESPONSIBLE FOR ATTENUATION IN VIRUSES OF THE ORDER DESIGNATED MONONEGAVIRALES</p> <p>(57) Abstract</p> <p>Isolated, recombinantly-generated, attenuated, nonsegmented, negative-sense, single stranded RNA viruses of the Order Mononegavirales having at least one attenuating mutation in the 3' genomic promoter region and having at least one attenuating mutation in the RNA polymerase gene are described. Vaccines are formulated comprising such viruses and a physiologically acceptable carrier. The vaccines are used for immunizing an individual to induce protection against a nonsegmented, negative-sense, single stranded RNA virus of the Order Mononegavirales.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 97/16718

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC 6 C12N15/45 C12N15/47 C12N15/40 C12N7/04 C07K14/12 C07K14/115 C07K14/135 A61K39/165 A61K39/155 C07K14/145 C07K14/08		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) IPC 6 C12N C07K A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 702 085 A (AKZO NOBEL NV) 20 March 1996 cited in the application  see the whole document	1-9, 17-21, 24-28, 31-37, 40-44
Y	EP 0 540 135 A (THE KITASATO INSTITUTE) 5 May 1993  see the whole document, especially SEQ.ID.1	1-6, 17-19, 24-26, 31, 33-35, 40-42
--- -/--		
<div style="display: flex; justify-content: space-between;"> <span><input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.</span> <span><input checked="" type="checkbox"/> Patent family members are listed in annex.</span> </div>		
* Special categories of cited documents : <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p>*A* document defining the general state of the art which is not considered to be of particular relevance</p> <p>*E* earlier document but published on or after the international filing date</p> <p>*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>*O* document referring to an oral disclosure, use, exhibition or other means</p> <p>*P* document published prior to the international filing date but later than the priority date claimed</p> </div> <div style="width: 45%;"> <p>*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>*Z* document member of the same patent family</p> </div> </div>		
Date of the actual completion of the international search  <div style="text-align: center; font-size: 1.2em;">29 May 1998</div>		Date of mailing of the international search report  <div style="text-align: center; font-size: 1.2em;">29 -06- 1998</div>
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer  <div style="text-align: center; font-size: 1.2em;">Mandl, B</div>

Form PCT/ISA/210 (second sheet) (July 1992)

## INTERNATIONAL SEARCH REPORT

Internat. Application No  
PCT/US 97/16718

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	STOKES A. ET AL.: "The complete nucleotide sequence of two cold-adapted, temperature-sensitive attenuated mutant vaccine viruses(cp12 and cp45) derived from the JS strain of human parainfluenza virus type 3 (PIV3)." VIRUS RESEARCH, vol. 30, 1993, pages 43-52, XP002051711 cited in the application see the whole document, especially Fig. 2 ---	1-3,7-9, 17,20, 21,24, 27,28, 32,33, 36,37, 40,43,44
A	WO 93 21306 A (THE MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK) 28 October 1993 see the whole document ---	1-9, 11-14, 17-46
A	EP 0 440 219 A (SCHWEIZ SERUM & IMPFINST) 7 August 1991  see the whole document, especially examples ---	1-9, 11-14, 17-46
A	RADECKE F. ET AL.: "RESCUE OF MEASLES VIRUSES FROM CLONED DNA" EMBO JOURNAL, vol. 14, no. 23, 1 December 1995, pages 5773-5784, XP002022952 cited in the application see the whole document ---	1-6, 17-19, 24-26, 31, 33-35, 40-42
A	MORI T. ET AL.: "Molecular cloning and complete nucleotide sequence of genomic RNA of the AIK-C strain of attenuated measles virus." VIRUS GENES, vol. 7, no. 1, 1993, pages 67-81, XP002051752 see the whole document, especially Figure 2 and Table 3 ---	1-6, 17-19, 24-26, 31, 33-35, 40-42
A	CROWE J.E. ET AL.: "Acquisition of the ts phenotype by a chemically mutagenized cold-passaged human respiratory syncytial virus vaccine candidate results from the acquisition of a single mutation in the polymerase (L) gene." VIRUS GENES, vol. 13, no. 3, February 1996, pages 269-273, XP002066591 see the whole document ---	11-14, 22,23, 29,30, 38,39, 45,46
	---	

-/--

## INTERNATIONAL SEARCH REPORT

Internati Application No  
PCT/US 97/16718

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	COLLINS P.L. ET AL.: "Production of infectious human respiratory syncytial virus from cloned cDNA confirms an essential role for the transcription elongation factor from the 5' proximal open reading frame of the M2 mRNA in gene expression and provides a capability for vaccine development." PROC. NATL. ACAD. SCI. USA, vol. 92, December 1995, pages 11563-11567, XP002066592 cited in the application see the whole document ---	11-14, 22,23, 29,30, 38,39, 45,46
A	WO 93 21310 A (AMERICAN HOME PROD ;US (US)) 28 October 1993  see the whole document ---	11-14, 22,23, 29,30, 38,39, 45,46
A	EP 0 567 100 A (AMERICAN CYANAMID CO) 27 October 1993  see page 13, line 9 - line 53 -----	11-14, 22,23, 29,30, 38,39, 45,46

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 97/16718

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
  
Remark : Although claims 25-30 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:  
  
(inventions 1,2,4) 4-9,11-14,18-23,25-32,34-39,41-46  
(completely) and 1-3,17,24,33,40 (partially)
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☒ No protest accompanied the payment of additional search fees.



## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 4-6,18,19,25,26,31,34,35,41,42 (complete) and 1-3, 17,24,33,40 (partially)

Recombinantly generated, attenuated Morbillivirus having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; nucleic acid molecules encoding said virus; composition comprising said nucleic acid together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

2. Claims: 7-9,20,21,27,28,32,36,37,43,44 (complete) and 1-3, 17,24,33,40 (partially)

Recombinantly generated, attenuated Paramyxovirus having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; nucleic acid molecules encoding said virus; composition comprising said nucleic acid together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

3. Claims: 10 (complete) and 1-3,17,24,33,40 (partially)

Recombinantly generated, attenuated Rubulavirus having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; composition comprising a nucleic acid molecule encoding said virus together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

4. Claims: 11-14,22,23,29,30,38,39,45,46 (complete) and 1,2, 17,24,33,40 (partially)

Recombinantly generated, attenuated Pneumovirinae having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; composition comprising a nucleic acid molecule encoding said virus together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 15 (complete) and 1,17,24,33,40 (partially)

Recombinantly generated, attenuated Rhabdoviridae having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; composition comprising a nucleic acid molecule encoding said virus together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

6. Claims: 16 (complete) and 1,17,24,33,40 (partially)

Recombinantly generated, attenuated Filoviridae having at least one attenuating mutation in the 3' genomic promoter region and at least one attenuating mutation in the RNA polymerase gene; vaccine comprising it; composition comprising a nucleic acid molecule encoding said virus together with an expression vector providing all trans-acting proteins necessary for the production of said attenuated virus and a method of producing it.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No  
PCT/US 97/16718

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0702085 A	20-03-1996	CA 2154023 A CN 1121959 A JP 8168381 A	19-01-1996 08-05-1996 02-07-1996
EP 540135 A	05-05-1993	JP 5103665 A KR 9512759 B US 5654136 A	27-04-1993 21-10-1995 05-08-1997
WO 9321306 A	28-10-1993	AU 687738 B AU 4289393 A CN 1082606 A EP 0636172 A ZA 9302588 A	05-03-1998 18-11-1993 23-02-1994 01-02-1995 02-11-1993
EP 0440219 A	07-08-1991	AU 7007491 A CA 2035386 A JP 4211377 A	08-08-1991 03-08-1991 03-08-1992
WO 9321310 A	28-10-1993	AU 683840 B AU 4290793 A CA 2118509 A EP 0640128 A HU 69992 A IL 105456 A JP 7505891 T MX 9302278 A NZ 252857 A PL 172087 B ZA 9302766 A	27-11-1997 18-11-1993 28-10-1993 01-03-1995 28-09-1995 05-12-1996 29-06-1995 29-07-1994 24-06-1997 31-07-1997 20-10-1994
EP 0567100 A	27-10-1993	AU 3705793 A CA 2094464 A JP 6022756 A NZ 247444 A ZA 9302763 A	28-10-1993 22-10-1993 01-02-1994 22-12-1994 08-02-1994

